

April 27, 2001
Version 2.1
Task Order Number CSC-01-11

生物软件网翻译制作

翻译人：charlson & zhanglinghui

PatentIn 3.1 使用手册

批准人:

项目经理 日期

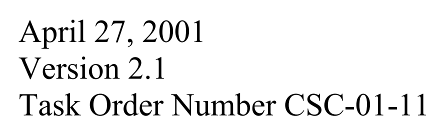
原作者:

Computer Sciences Corporation
2611 Jefferson Davis Highway, Suite 10,000
Arlington, VA 22202-4016



档案编号与说明	
PatentIn 3.1 使用手册	PTO-OP-07
USPTO 合同 **50-PBPT-7-00004**	系统开发和维护
任务名称	PatentIn 3.1
任务类型	不连续
任务编号	CSC-01-11





版本更新历史

版本号	变更日期	变更原因	变动的页码	批准人签名
N/A	11/10/98	Lockheed Martin Corporation 制作的原始文档	所有页码	J. Paccassi
1.0 版草稿	06/03/99	CSC 制作的原始文档	所有页码	J. Paccassi
1.0.版	09/16/99	整合了 USPTO 的注释和来自 FQT 的校正	i, ii, 1v, 3-5, 3-6, 4-2, 4-3, 4-5, 4-6, 5-2, 5-3, 5-5, 6-2, 6-4, 6-7, 7-2, 7-3, 8-3 through 8-9	J. Paccassi
1.1 版	06/02/00	更新了从 JAVA 变为使用 C++导致的文档变化	所有页码	J. Paccassi
1.2 版	06/22/00	USPTO 整合了来自 FQT 的校正	3-5, 3-10, 5-4, 5-10, 6-2, 8-4	J. Paccassi
2.0 版	01/26/01	PatentIn 3.1 版发布	所有页码	J. Paccassi
2.1 版	04/27/01	PatentIn 3.1 发布版的更新	所有页码	

April 27, 2001
Version 2.1
Task Order Number CSC-01-11

任务执行概述

本使用手册的内容与 PTO-OP-07 手册的内容并不一致。反而，根据 WPI Task Management Plan (TM02) 2.4.4.4 部分的要求，只能根据最初准备的文档的内容、格式进行更新。

本使用手册根据 PTO 的要求进行更新，任务号：CSC-01-11。

目录

目录(续)

目录(续)

表格目录

图目录

第一部分

介绍

第一部分
介绍

1.1 目的

PatentIn 可以加快建立专利申请中包含的序列列表，可以利用它输入序列、验证数据、建立序列列表文件，并能打印或保存到其他移动介质中以便提交申请。本手册介绍如何使用 PatentIn。

1.2 惯例

本手册使用了一目了然的符号提示和标准按键操作，使用的惯例见下表 1-1。

表 1-1: 文档中的惯例

符号	含义	举例
	利用鼠标点击	左键单击菜单中的 Open 命令
	按键盘上的按键	在编辑栏进行输入文字
粗体	文件名、菜单项或功能键名称	点击 **Exit**.
	对用户的提示	提示: 此栏必须填写。

1.3 概述

设计 PatentIn 的目的是加快向美国专利与商标局(USPTO)申请包含核酸或多肽序列的专利的过程。

PatentIn 遵循所有世界知识产权组织（World Intellectual Property Organization Standard，WIPO) ST.25 和相关的美国法规要求的文件格式，如“Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Disclosures”。程序可以运行于 Windows 95，Windows 98，Windows NT 和 Windows 2000，窗口显示为英语。因为 PatentIn 生成的序列文件符合 ST.25 的要求，因而适用于全球范围。

为保持易用性，程序设计中遵循标准的用户窗口。

PatentIn 中包括下列工具：

- 序列编辑器

序列编辑器是 PatentIn 中的基本工具，可以在其中输入、修改核酸或蛋白质序列，也可以从其他编辑器或 Word（前提是保存为 ASCII 文本文件）导入序列列表文档。

使用 PatentIn 时，可以随时输入任意顺序的数据，也可以添加、删除、修改序列数据，还可以保存未完成的任务待以后完成。PatentIn 对任务是否存在或存在于什么机器或设备上没有要求，用户可以自由通过 email 发送到客户端或互相邮寄以便浏览或更新。

- 序列生成器

输入专利申请所需的所有数据后，PatentIn 就可以生成申请文件了。申请文件是计算机可读的，并符合 ST.25 规范的序列列表文件。

April 27, 2001
Version 2.1
Task Order Number CSC-01-11

第二部分
操作系统要求和运行 PATENTIN 3.1

第二部分
操作系统要求和运行 PATENTIN 3.1

2.1 操作系统要求

PatentIn 是可以独立运行的程序，可以从美国专利与商标局(USPTO)的网站上下载，运行于 Windows 95/98/NT/2000 系统中。推荐至少有 64 MB 内存，对于巨大的专利申任务需要更多的内存，对于象包含 100,000 条序列或一条序列达到 1M 大小的申请，至少需要 128 MB 内存。PatentIn 3.1 需要的硬盘空间是 2.5 MB，另外需要额外的空间以保存任务文件和序列列表文件。如果想使用导入 PatentIn 2.1 版任务的功能，还需要安装 Micrsoft® Access 及其部分扩展功能（如只是导入序列文件，则没有此要求）。

要使 PatentIn 正常工作，环境变量 “TMP”要被设置为指向一个已经存在的目录，而环境变量“PATH”路径中也必须设置为包含 DOS 备份命令，不过多数情况下在 Windows 中安装完成后都符合这些要求。

ⓘ对巨长序列或大量序列用户的提示： USPTO 找到一个非常适用于非常大的文本文件的程序，其安装后 60 天内有效的试用版可以从 www.fileviewer.com <http://www.fileviewer.com>下载。程序名称为"V"， 版本号为 2000 SR-1。PTO 用 60MB 和 120MB 大的文件进行了测试，结果 “非常棒”， 也应该可以处理任意大小的文件。PTO 是在膝上型计算机、Win98 环境下进行测试的（安装时可能需要管理员权限）。

2.2 运行 PATENTIN

PatentIn 为单台计算机安装所设计，需要的话可以下载到个人电脑上， 而任务文件可以在其他地方保存。程序可以从 USPTO 网页上下载，地址是: ftp://ftp.uspto.gov/pub/patentin/patentin3116.exe。点击链接即可以下载 PatentIn 3.1 然后安装到计算机上。完成安装后，桌面会出现一个 PatentIn 3.1 图标， 鼠标左键双击图标，就可以运行 PatentIn 3.1。

第三部分
启动程序

第三部分
启动程序

3.1 序列窗口

鼠标双击 PatentIn 3.1 程序图标，出现序列输入窗口（**Sequence Screen**，图 3-1）。

序列窗口 （**Sequence Screen**，图 3-1） 是程序主窗口。

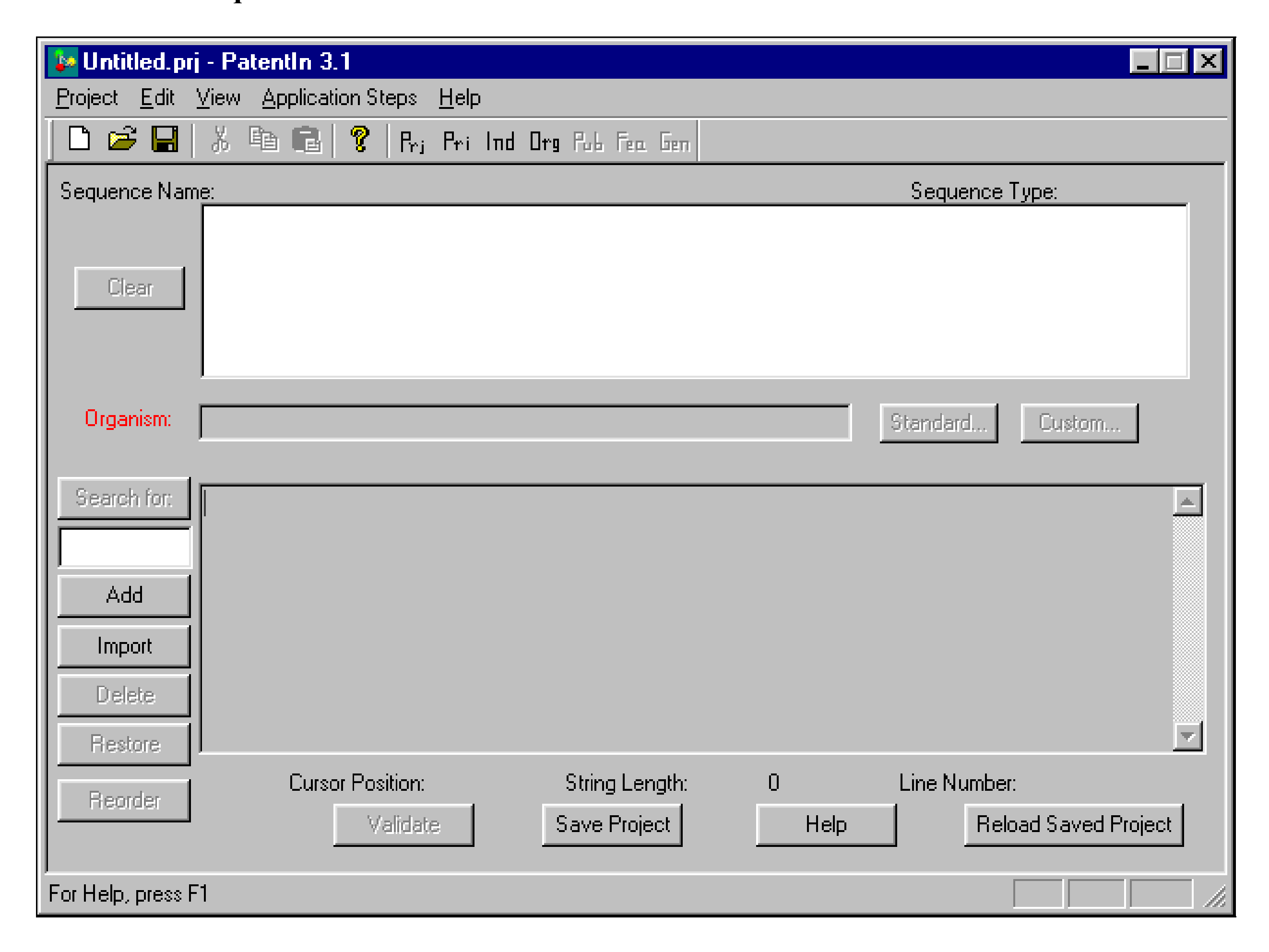


图 3-1: Sequence Screen

序列窗口（图 3-1）中共 5 个下拉菜单，其中三个可以随时与系统交互作用：Project，Application Steps 和 Help。另外两个菜单即 Edit 和 View，是微软 Microsoft®操作系统类型的菜单。开始一个任务后可以选择三个菜单中的任一个，PatentIn 启动时开始一个名称为"Untitled"的空的任务，也可以通过任务菜单（**Project Menu**，图 3-2）打开一个已经存在的任务。

3.2 任务菜单

可以在任务菜单（**Project Menu**，图 3-2） 中创建或保存一个任务。选择 **Save**，出现 **"**另存为**"**窗口（**Save As Screen**，图 3-3），可以给任务命名并保存。选择 **"Open"**，出现打开窗口（**Open Screen**，图 3-4），选择先前保存的项目，然后打开。" Exit PatentIn 3.1"选项用来关闭窗口。菜单项是灰色的，除非有一个打开的任务或是有一个输出的文件。

下图是任务菜单（**Project Menu**，图 3-2）选项，可以看到任务菜单中有一列可选的子菜单选项。任务菜单和当前活动的任务名称在窗口的左上角。这个例子中，出现的是 Untitled，是因为并没有打开或保存一个任务。

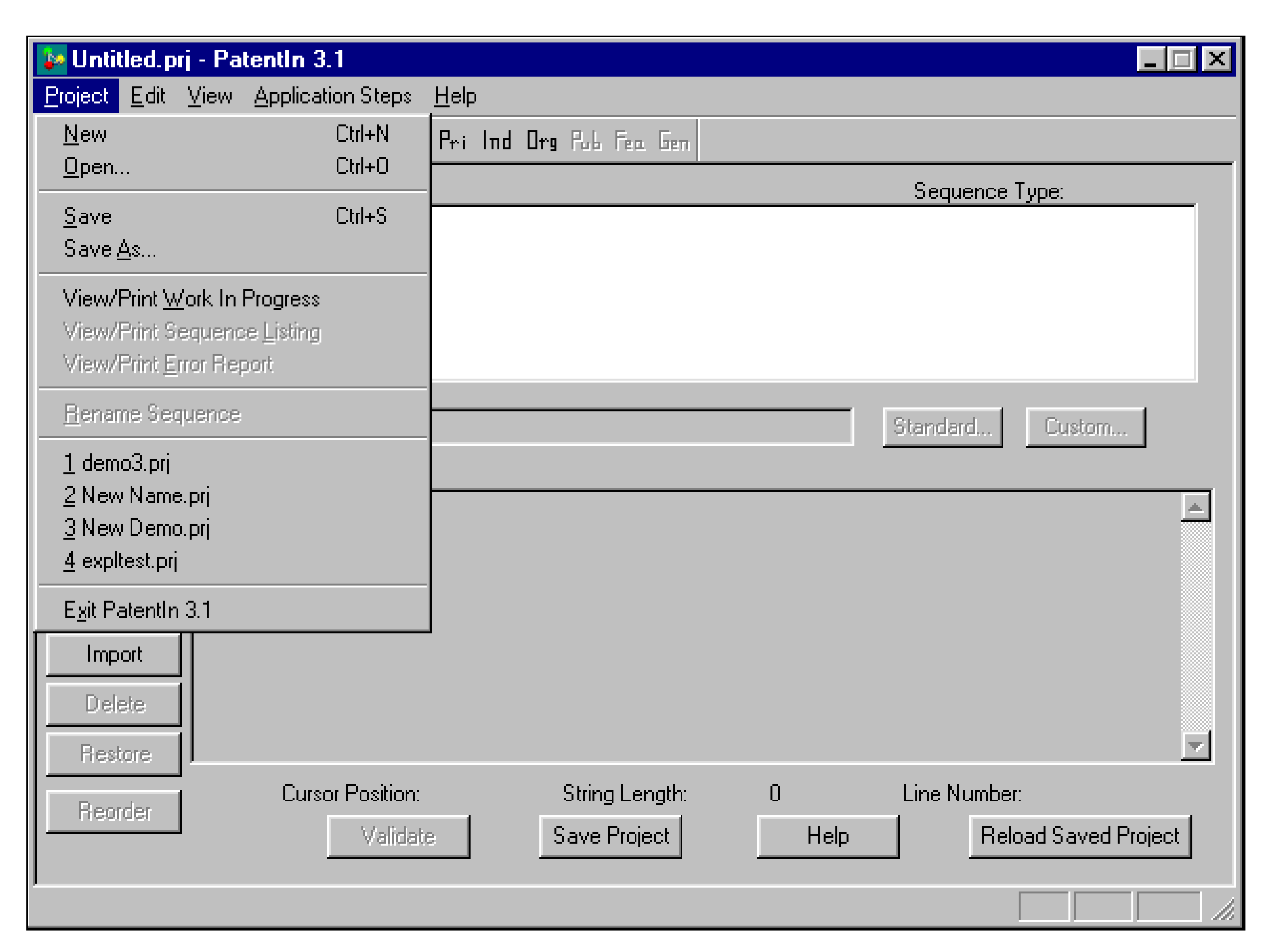


图 3-2: Project Menu

3.3 建立和保存新任务

建立或保存新任务：

打开主窗口即序列窗口即可建立新文件或者:

1. 选择 **Project** 菜单中的 **New**，可以清除当前任务中的所有信息；
2. 选择 **Project** 菜单中的 **Save AS**，出现Save As Screen (**Save As Screen**，图 3-3)；
3. 在 **File Name** 对话框中输入新文件名称；
4. 点击 **Save** 建立新文件。
5. 新文件的名称出现在屏幕的左上角。

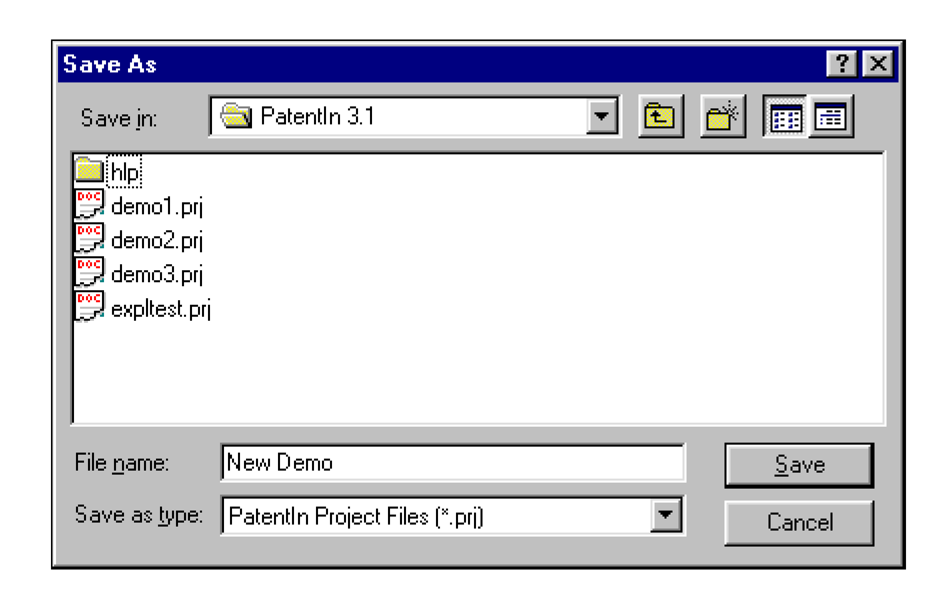


图 3-3: Save As Screen

提示： **Save** 和 **Save As** 只保存任务文件（.prj 格式）。在生成的任务中所包含的列表以纯文本格式保存，其名称和任务的名称相同，不过扩展名为 “ST25.txt”。

3.4 打开任务

打开现成的任务：

1. 选择 **Project** 菜单中的 **Open**，出现打开对话框（**Open Screen**，图 3-4）；
2. 打开文件所在的目录；
3. 双击文件名称，打开文件；
4. 回到程序主窗口中，此时打开的任务名称出现在屏幕的左上角，表示任务处于活动状态。

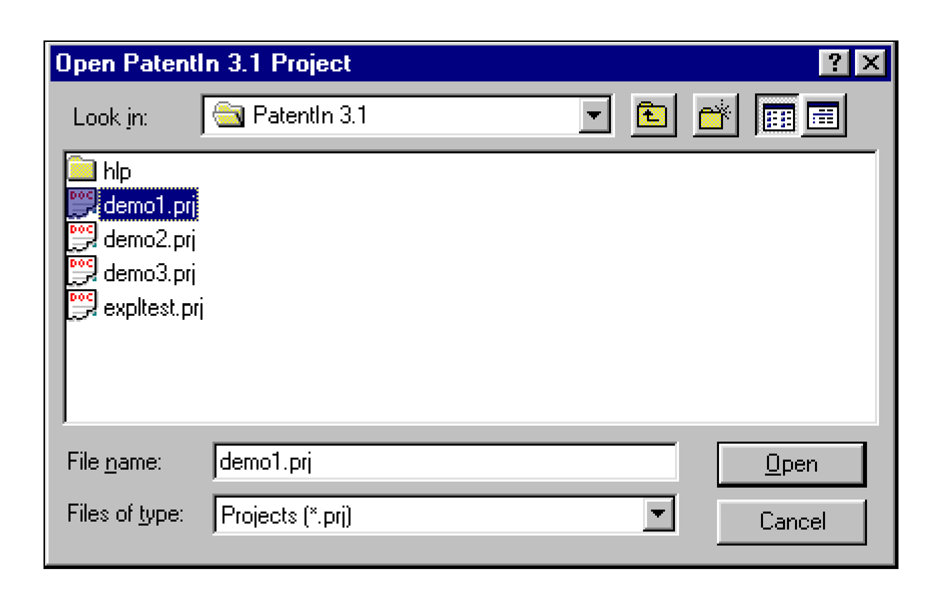


图 3-4: Open Screen

ⓘ 对处理巨大序列和/或巨大数量序列的用户的提示：处理特长序列和/或非常多的序列时：从内存中清除一个很大任务时需要一些时间，这一点在立即重新打开任务时明显。

3.5 保存任务

保存一个任务：

第一次保存任务时，程序会提醒用户输入文件的名称。

1. 选择 **Project** 菜单中的 **Save**，如果此前任务没有被命名，会出现保存窗口（**Save As Screen**，图 3-5）。否则，任务以先前打开或创建的名称保存；
2. 选择保存文件的目录；
3. 在 **File Name** 对话框中输入新文件名称；
4. 点击 **Save**，以新文件名保存任务；
5. 回到了 PatentIn 程序主窗口，新文件名出现在左上角，提示任务处于活动状态。

以不同的文件名称保存：

1. 选择 **Project** 菜单中的 **Save As**，出现“另存为”窗口（**Save As Screen**，图 3-5）；
2. 选择保存文件的目录；
3. 在 **File Name** 对话框中输入新文件名称；
4. 点击 **Save**，以新文件名保存任务。
5. 回到了 PatentIn 程序主窗口，新文件名出现在左上角，提示任务处于活动状态。

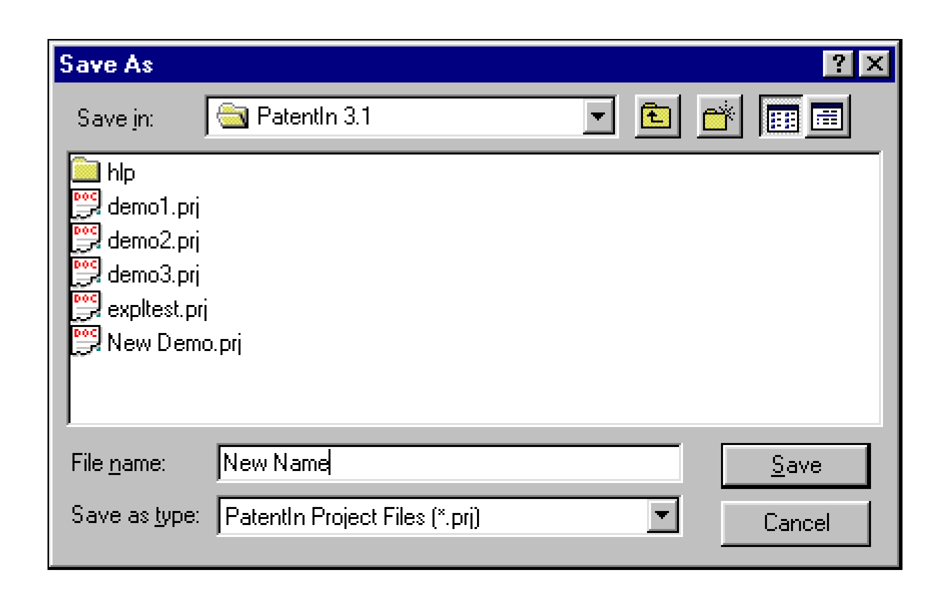


图 3-5: Save As Screen

3.6 浏览工作文件

用户可以通过创建一个工作文件来浏览当前工作进程，通过工作文件可以在一处观看整个任务的数据而不用浏览一个个窗口。注意不要把工作文件与序列列表混淆。

查看工作文件：

在 **Project** 菜单中，选择 **View/Print Work in Progress**。

提示：　通过选择 **Create Work File**，创建工作文件。.

3.7 查看工作进度

PatentIn 通过查看工作进度窗口（**View Work in Progress Window**，图 3-6）提供了全屏幕显示专利申请内容。

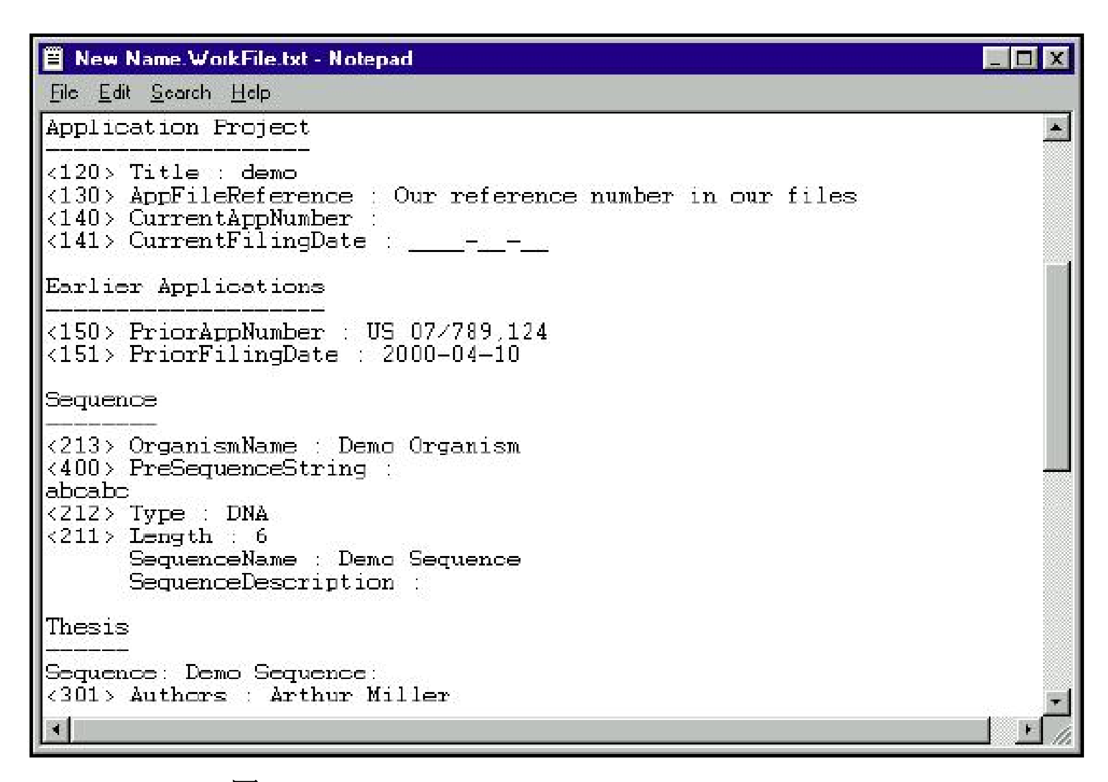


图 3-6: View Work in Progress Window

查看当前专利申请的内容：

1. 在 **Project** 菜单中，选择 **View Work in Progress** 项；
2. 点击一个文件，然后点击 Print，可以打印这个文件的出错报告；

3. 点击 File Exit，退出程序窗口。

3.8 查看序列列表

PatentIn 提供了序列列表查看窗口（**View Sequence Listing Window**，图 3-7）观看屏幕显示的序列列表。

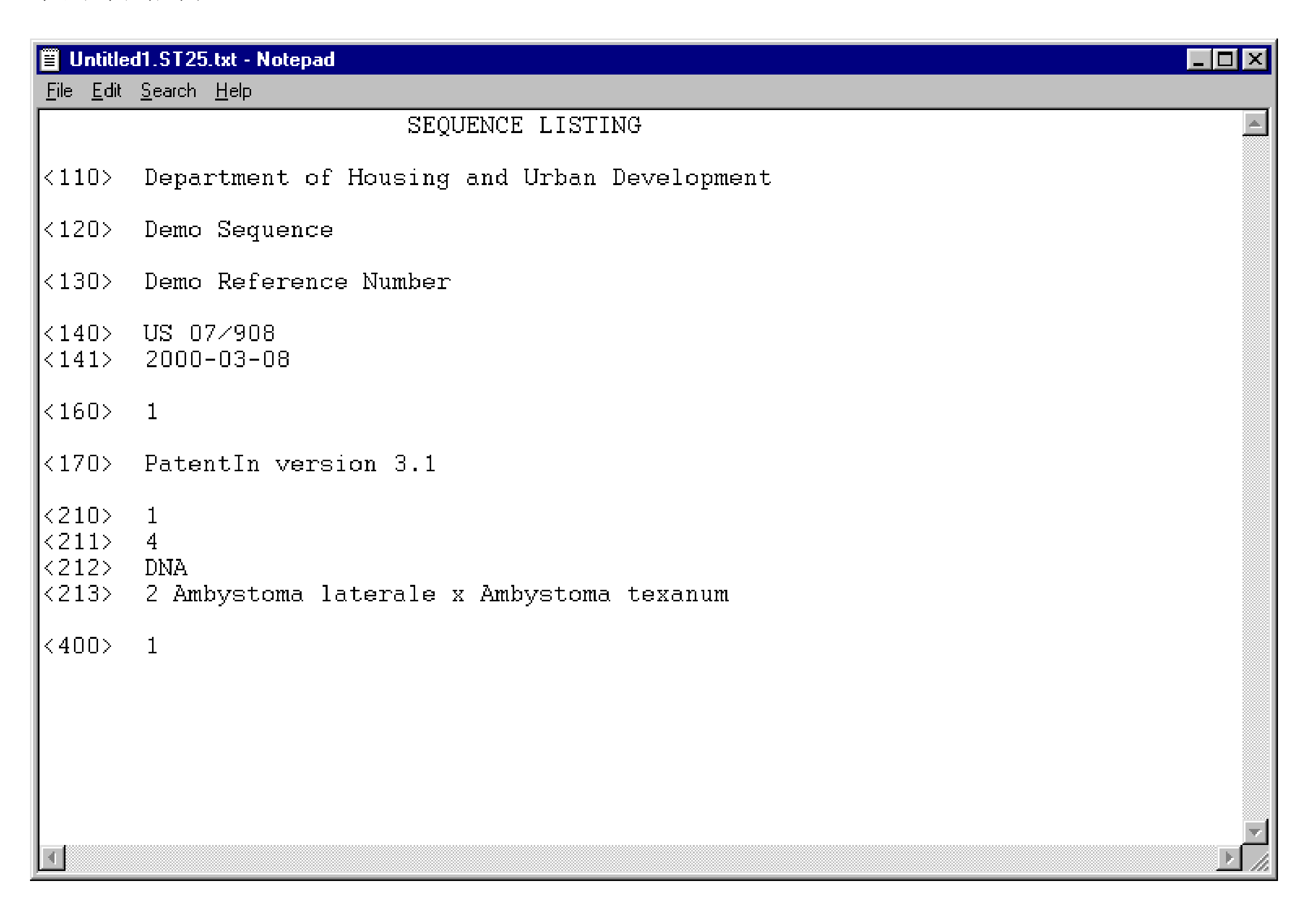


图 3-7: View Sequence Listing Window

观看序列列表的方法：

1. 在 **Project** 菜单中， 选择 **View Sequence Listing** 项；
2. 点击 **File** 中的 **Print**，可以打印文件出错报告；
3. 点击 **File** 中的 **Exit**，就退出窗口。

提示：　　必须首先产生序列。

3.9 浏览出错报告

通过出错报告查看窗口（**View Error Report Window**，图 3-8）可以观看显示在屏幕上的出错报告，显示了打开任务可能存在的错误。

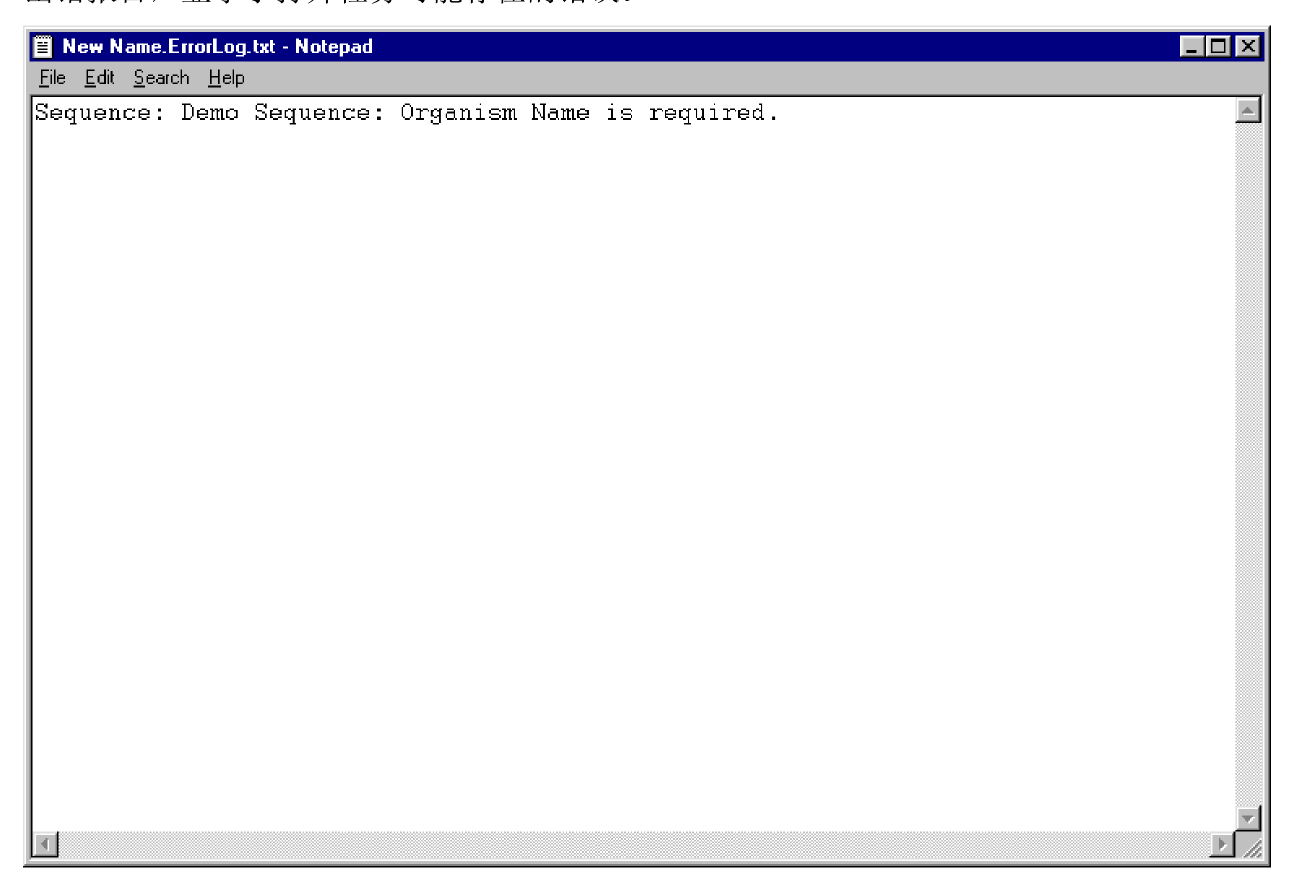


图 **3-8: View Error Report Window**

观看出错报告的方法：

1. 在 **Project** 中，选择 **View Error Report**；
2. 点击 **File** 中的 **Print**，打印出错报告；
3. 点击 **File** 中的 **Exit**，退出程序窗口。

3.10 序列重新命名

PatentIn 的新特点是能够改变一条序列的名称（**Rename Sequence Screen**，图 3-9）。

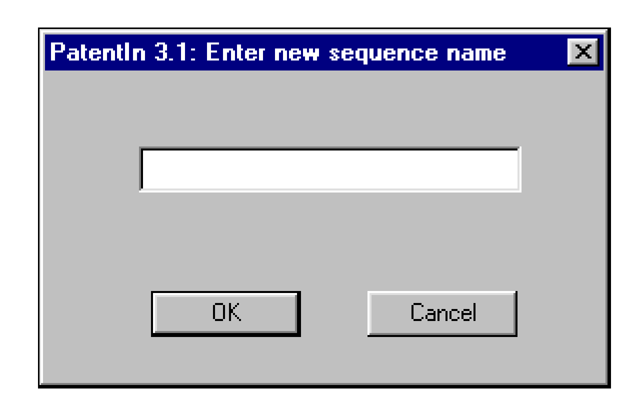


图 3-9: Rename Sequence Screen

1. 点击序列的名称，打开 **Rename Sequence** 窗口；
2. 选择 **Project** 菜单中的 **Rename Sequence** 项；
3. 在 **Rename Sequence** 对话框中输入序列的新名称；
4. 点击 **OK** 按钮。

3.11 退出 PATENTIN

退出 PatentIn 时，如果任务没有保存，会提示用户询问是否保存当前任务（**Exit PatentIn Screen**，图 3-10）。

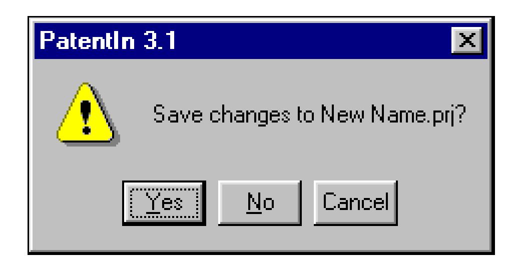


图 3-10: Exit PatentIn Screen

3.12 使用在线帮助

在线帮助对多数 PatentIn 窗口有效，通过按 F1 键或 help 按钮，出现帮助窗口，下面的例子是一个帮助窗口（**Help Screen**，图 3-11）。

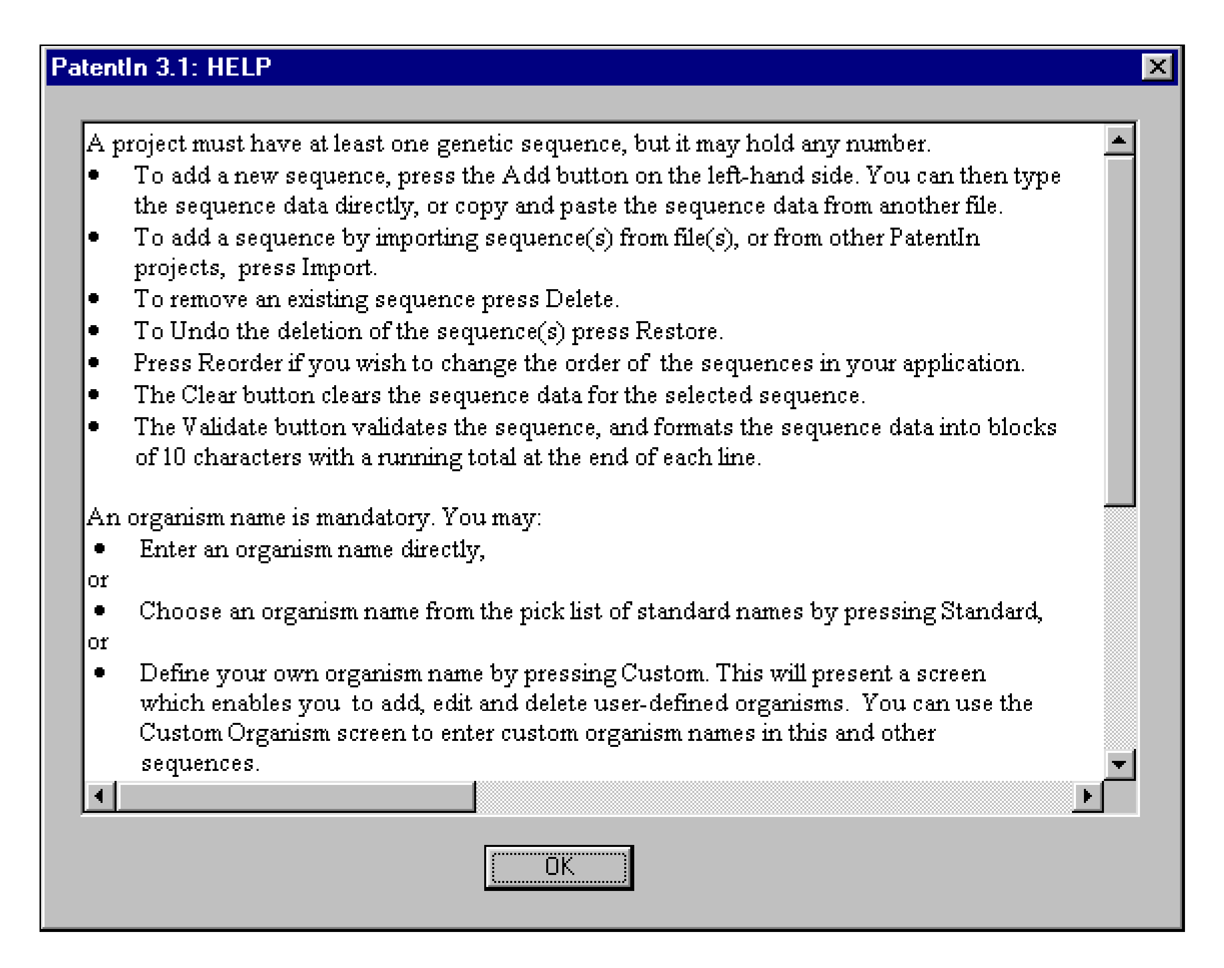


图 **3-11: Help Screen**

1. 点击 **OK** 按钮，即退出帮助窗口。

3.13 消息对话框

按下了一个动作按钮（如第四部分的 **Add**，**Project** 以及 **Application Data**），而窗口输入栏又没有输入内容时，会出现消息对话框窗口（**Message Dialog Screen**，图 3-12）。

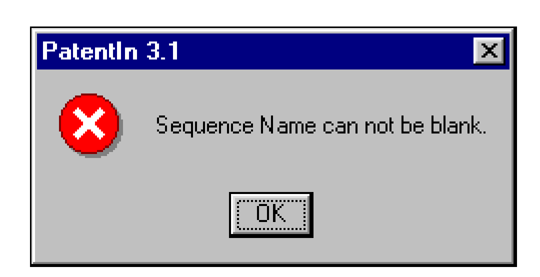


图 **3-12: Message Dialog Screen**

April 27, 2001
Version 2.1
Task Order Number CSC-01-11

第四部分
任务和专利申请的数据

第四部分
任务和申请的数据

创建序列列表后，可以继续向申请中添加其他信息。

4.1 申请步骤菜单

开始或建立或选择一个任务时，**申请步骤菜单**（**Application Steps Menu**，图 4-1）处于活动状态，任务名显示在窗口的左上角。下面的例子是打开一个现成的文件，任务名称为“New Name”。

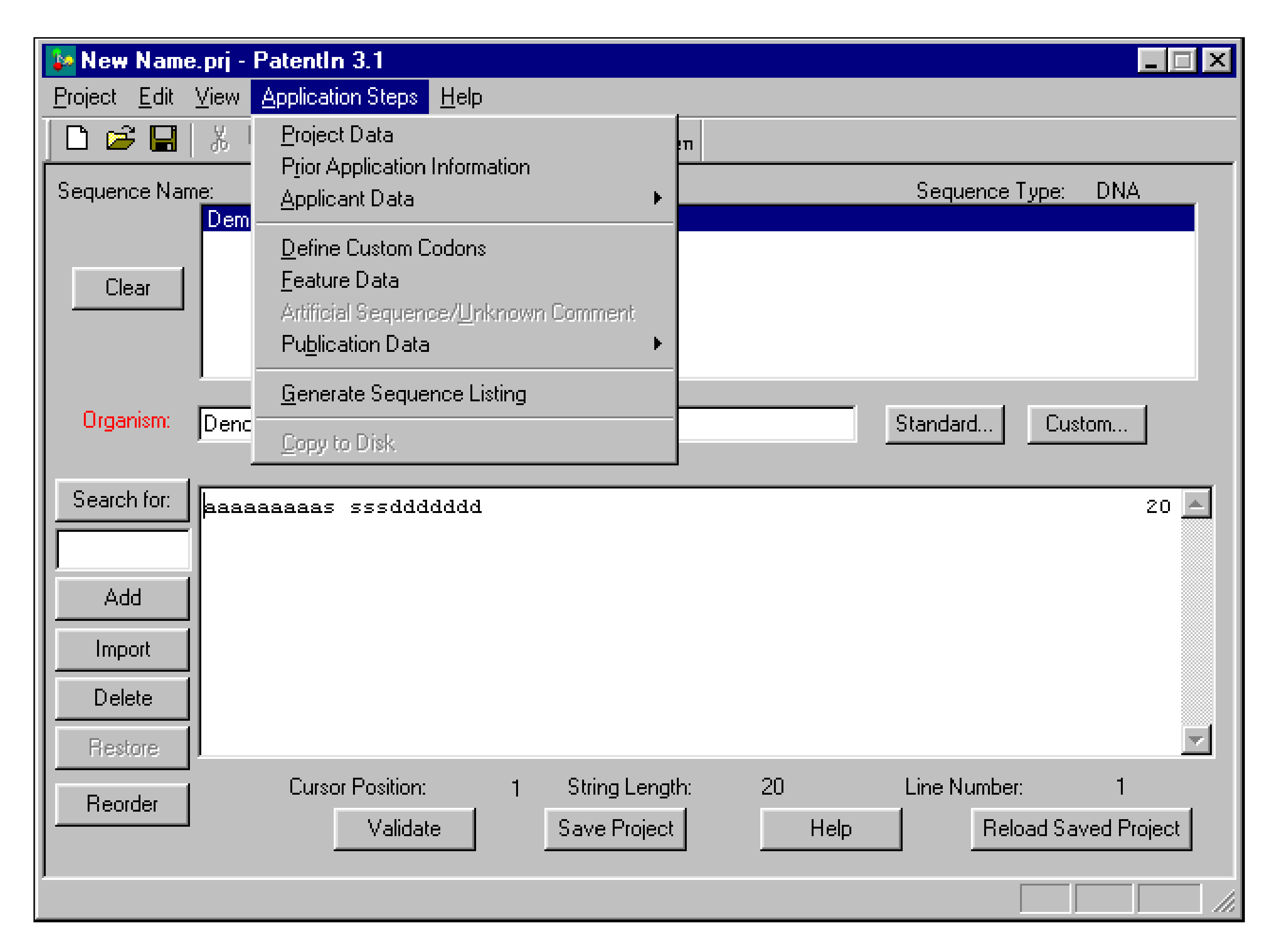


图 **4-1: Application Steps Menu**

4.2 任务数据

用户可以在任务数据窗口（**Project Data Screen**，图 4-2）的输入栏输入新发明的相关内容。这些信息对确定发明的主题和备案日期很主要。

ⓘ提示： 红色表示必须填写的地方（发明的名称及申请文件的参考文献）.

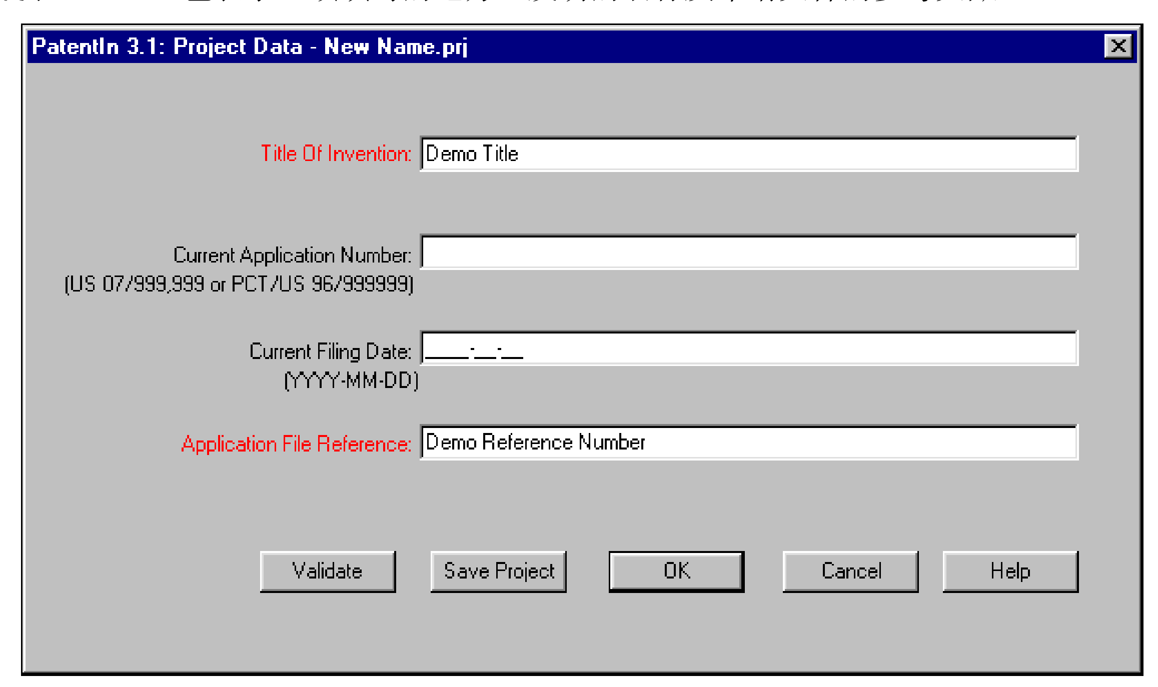


图 **4-2: Project Data Screen**

输入任务数据：

1. 🖮输入发明的名称（**Title of the Invention**），这是必填的内容；
2. 如果存在的话，🖮 输入当前申请号（**Current Application Number**）。一旦输入申请，则当前记录日期（**Current Filing Date**）必须填写；
3. 🖮输入当前纪录日期（**Current Filing Date**），格式是数字的：如 1989-06-28；
4. 🖮输入申请文件的参考文献（**Application File Reference**）；
5. 🖱点击 **Validate**，确认输入的信息；
6. 🖱点击 **Save Project**，保存输入的信息。

4.3 优先申请信息

优先申请的信息属于可选的输入内容，因为专利审查人员可以从申请文件库得到这些信息。优先的申请都包括在优先申请信息窗口中（**Prior Application Information Screen**，图 4-3），按输入的顺序显示在表中，可选中进行编辑或删除。

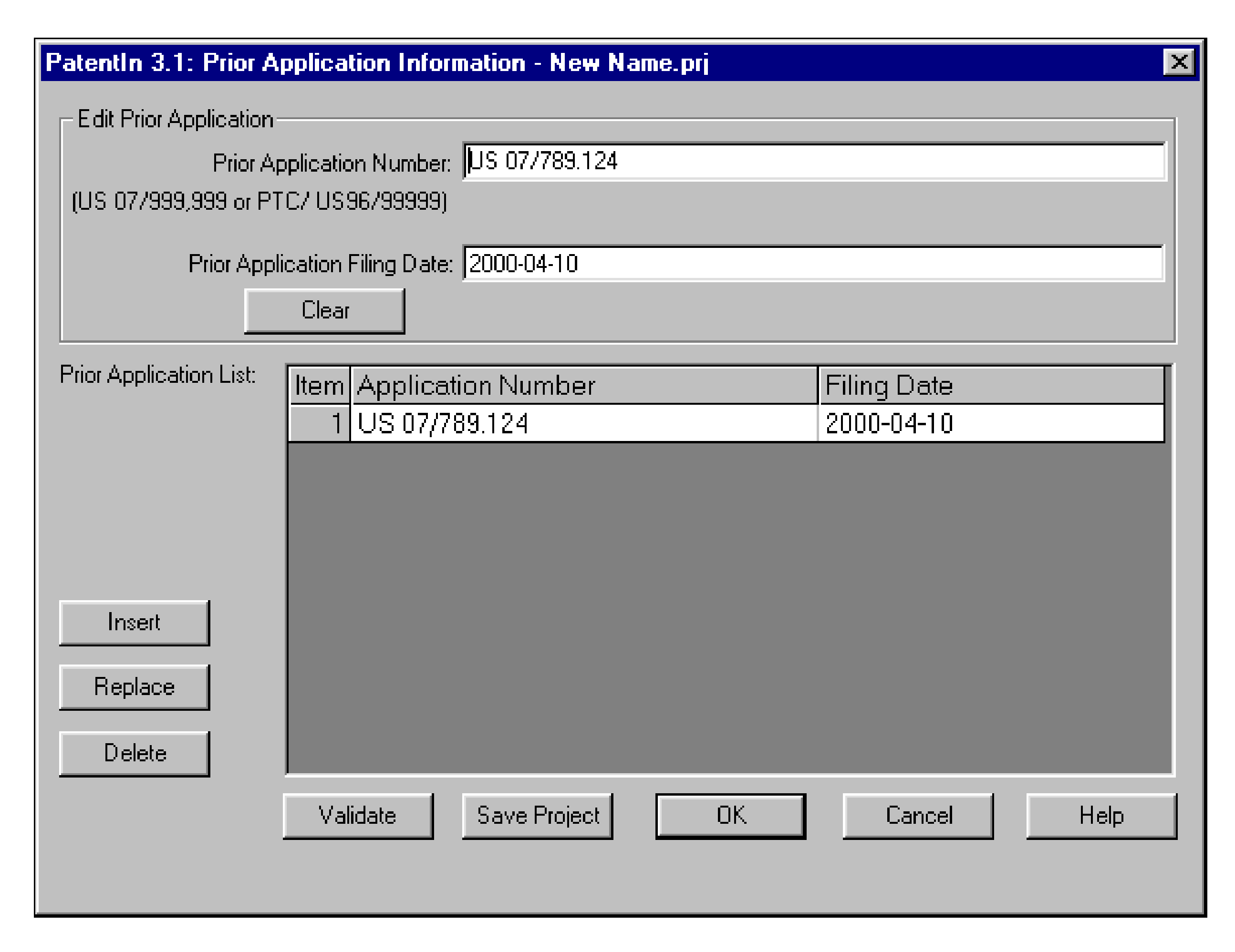


图 **4-3: Prior Application Information Screen**

输入优先申请的信息：

1. 输入优先申请号（**Prior Application Number**）。如果输入了优先申请号就必须填写优先申请的日期；
2. 输入优先申请纪录日期（**Prior Application Filing Date**），格式如：1989-06-28。
3. 点击 **Clear** 可以清除优先申请编辑（**Edit Prior Application**）区的信息；
4. 列表中插入信息的方法：选择在插入条目前面的信息，输入优先申请号（**Prior Application Number**）和优先申请纪录日期（**Prior Application Filing Date**），然后点击 **Insert** 按钮；
5. 替换列表中的条目：选择该条目，输入优先申请号（**Prior Application Number**）和优先申请纪录日期（**Prior Application Filing Date**），然后点击 **Replace** 按钮；
6. 删除条目：选择该条目，然后点击 **Delete** 按钮；
7. 点击 Validate，确认输入的信息。输入（或插入）到表中的数据就得到确认。编辑区的信息，没有插入，就得不到确认；

8. 🖱点击保存任务（Save Project）按钮，保存信息；
9. 🖱点击 OK 按钮，确认并关闭窗口。

ⓘ PATENTIN 的新特点：　　PATENTIN 3.1 中，点击 **OK** 后，如果编辑栏里的数据和选中的数据不同，将被插入到数据列表中。

4.4 申请数据

在序列窗口（**Application Steps Menu**，图 4-1）中可以输入个人或单位申请的信息，在 **Application Steps** 菜单里选择 **Applicant Data**，然后选择下一级菜单的个人（**Individual**）或单位（**Organization**）。如果选择了个人，出现个人申请窗口（**Individual Applicants Screen**，图 4-5），选择单位，出现单位申请窗口（**Organization Applicants Screen**，图 4-6）。

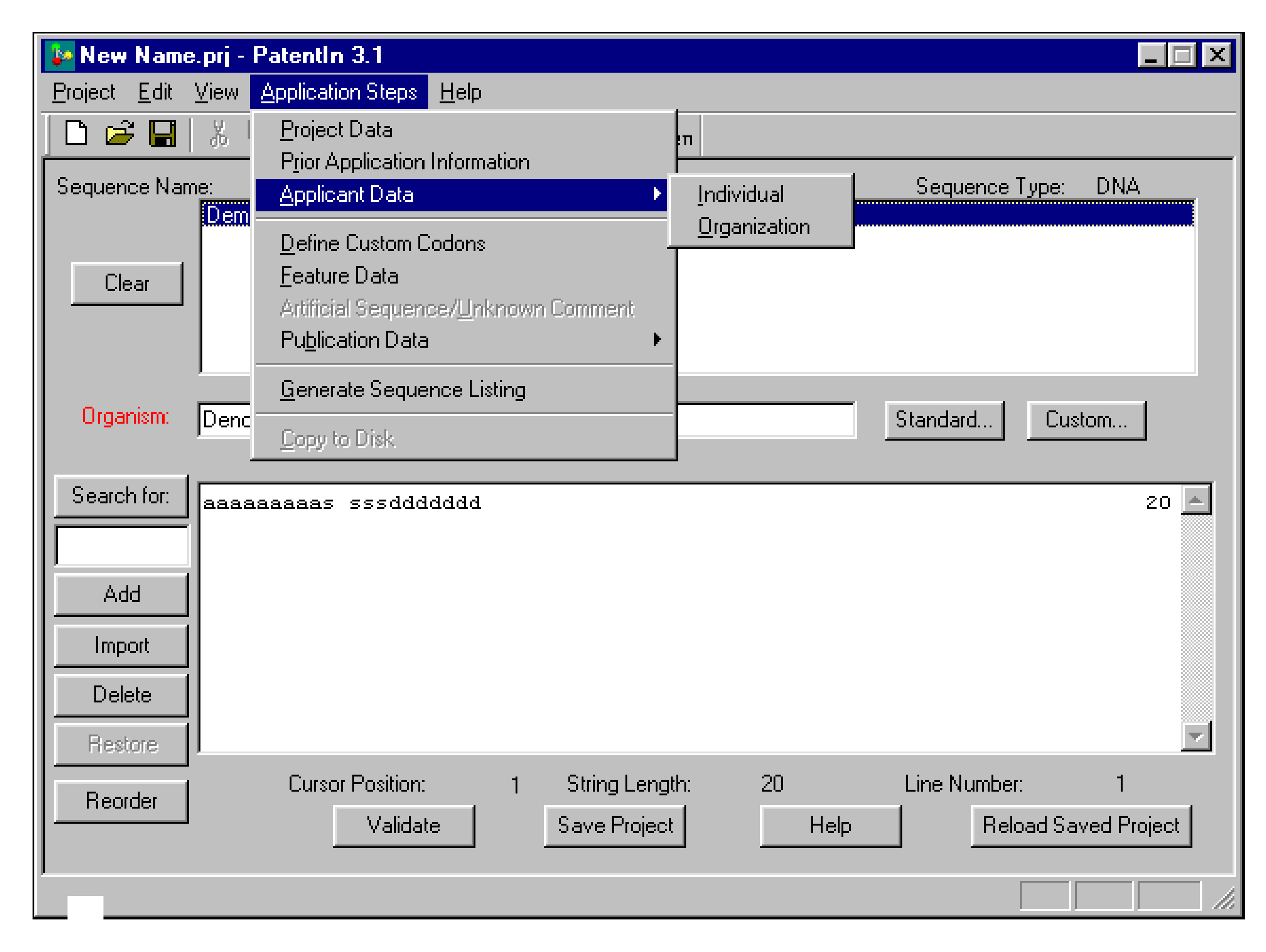


图 **4-4: Applicant Data Screen**

4.4.1 个人申请

个人申请窗口（**Individual Applicants Screen**，图 4-5）中可以输入申请人的信息，列表中只出现申请者的姓名，其他地方可由用户输入其他的信息。

ⓘ 提示： 红色的姓名栏（名-First Name，姓氏-Last Name）是必须填写的。

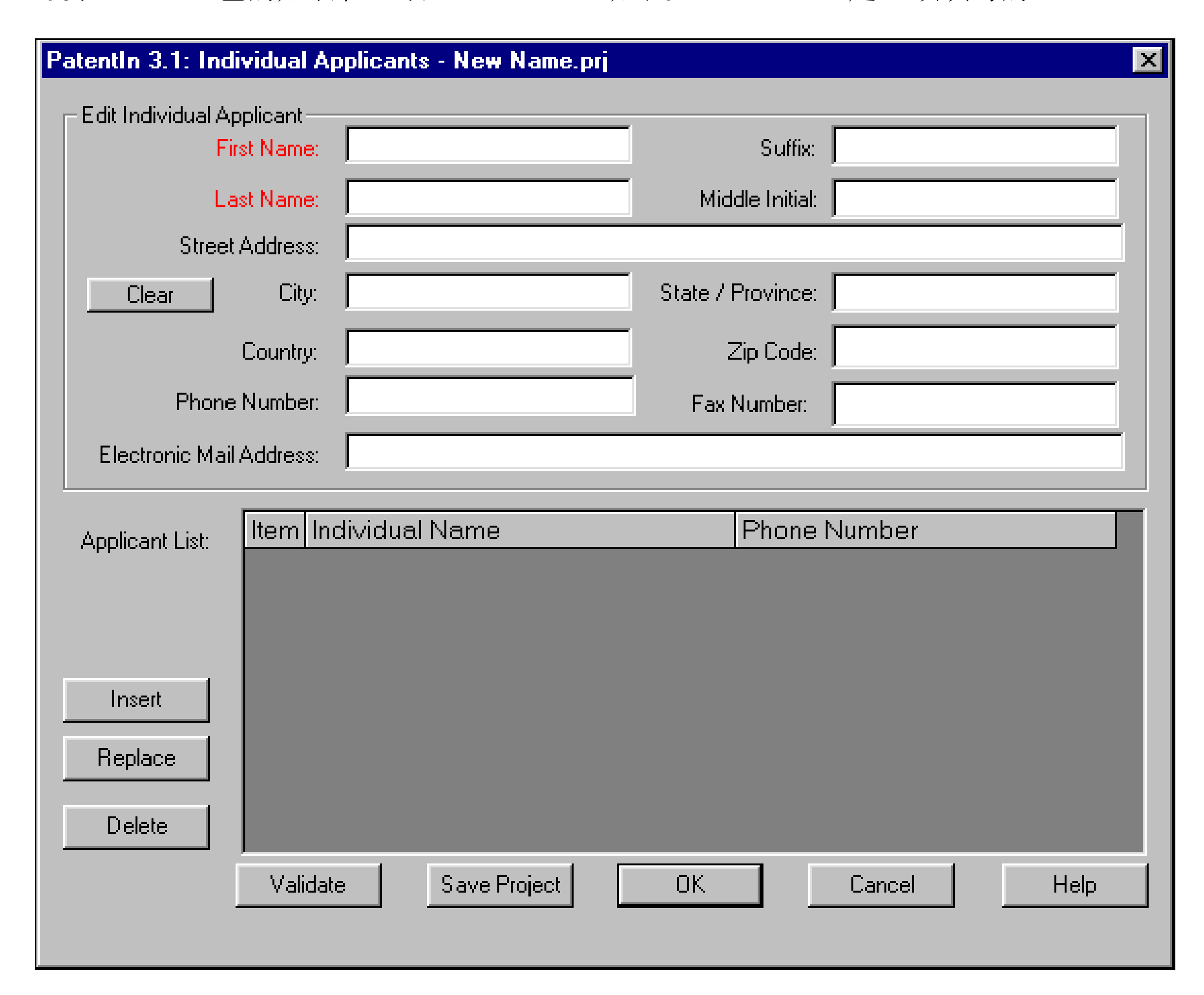


图 4-5: Individual Applicants Screen

输入申请者个人信息：

1. 在 **Application Steps** 中，🖰选择 **Applicant Data**，然后🖰 选择个人（**Individual**）；
2. 🖮输入申请者的姓氏（Last Name）；
3. 🖮输入申请者的名的前缀（Suffix）（如 Jr，III 等）；
4. 🖮输入申请者的名（First Name）；
5. 🖮输入申请者的中间名（Middle Initial）；

6. 输入街道地址，城市，州/省，国家，邮政编码，电话号码，传真号码以及电子邮件地址；
7. 点击 **Clear**，可以清除个人申请的信息；
8. 要插入信息，选择插入位置前的信息，输入 **Edit Individual Applicant**，然后点击 **Insert** 按钮；
9. 要替换某条信息，选中这条信息，输入 **Edit Individual Application**，然后点击 **Replace** 即可；
10. 要删除某条信息，选中这条信息，然后点击 **Delete** 即可；
11. 要确认输入的信息，点击 Validate，输入到表格中的数据立即被确认。但是编辑区的信息，由于没有插入，则不被确认；
12. 点击 OK 确认并关闭这个窗口；
13. 重复 2-12 步，就可以加入其他的信息。

提示：　　电话、传真号码，邮政编码不会再次确认。

PATENTIN 的新特点：　　在 PATENTIN 3.1 中，点击 **OK** 后，编辑区的数据如果和列表中的数据不同，则被加入列表中。

4.4.2 单位申请

单位申请窗口（**Organization Applicants Screen**，图 4-6）中可以输入申请单位的信息，象个人申请一样，序列列表中只有申请单位的名字，其他可输入区则可以由申请者输入其他信息。

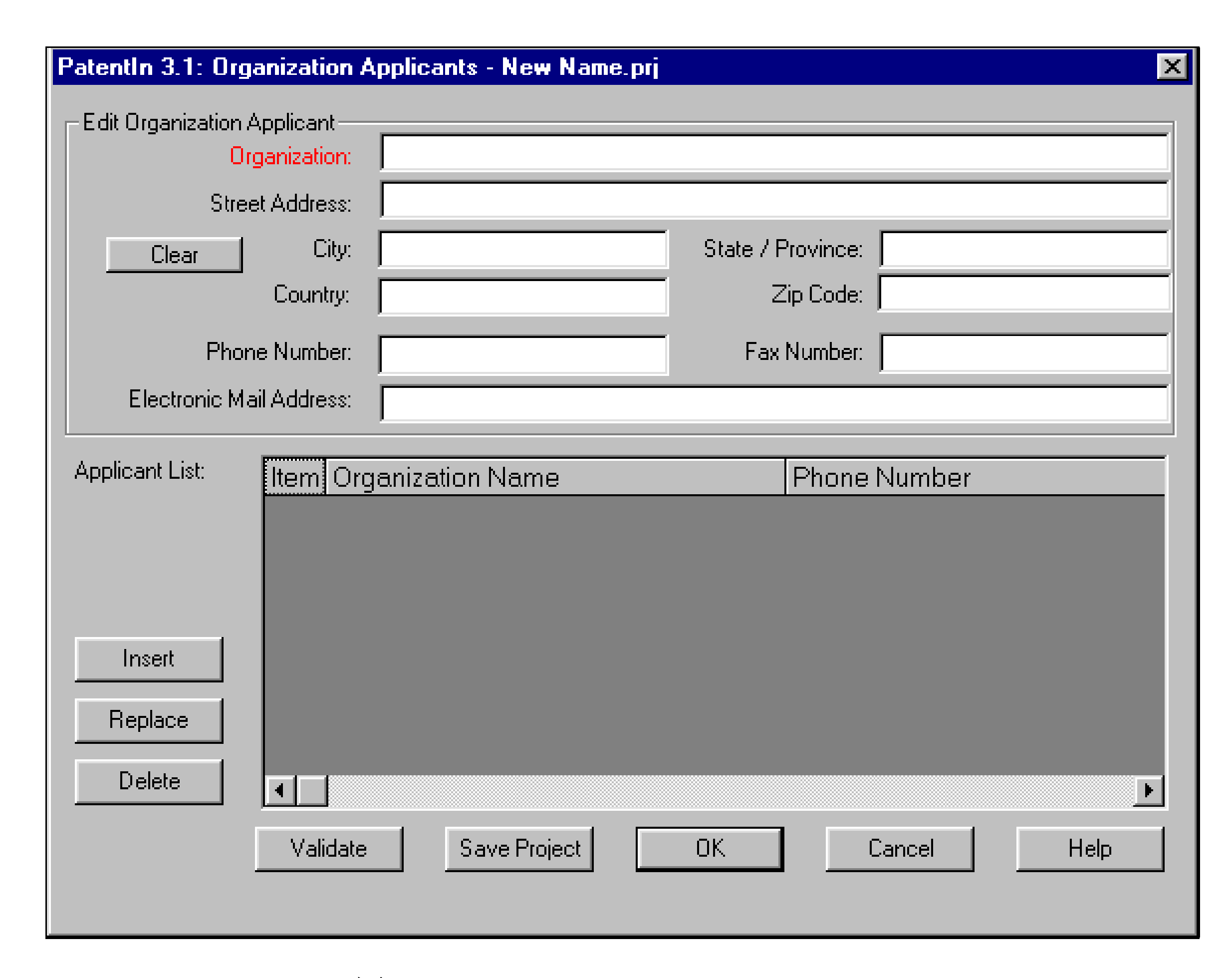


图 4-6: Organization Applicants Screen

输入申请者个人信息：

1. 在 **Applicant Steps** 菜单中， 选择 **Applicant Data** 子菜单， 然后选择单位（**Organization**）；
2. 输入单位（**Organization**）名称；
3. 然后输入单位的街道地址，城市，州/省，国家，邮政编码，电话号码，传真号码以及电子邮件地址；
4. 点击 Clear，可以清除窗口中 Edit Organization Application 栏中关于单位的信息；
5. 要插入信息，选择插入位置前的信息, 输入 **Edit Organization Applicant**，然后点击 Insert 按钮；
6. 要替换某条信息， 选中这条信息， 输入 **Edit Organization Application**，然后点击 **Replace** 即可；
7. 要删除某条信息，选中这条信息，然后点击 **Delete** 即可；
8. 要确认输入的信息，点击 Validate，输入（插入）到表格中的数据立即被确认。但是编辑区的信息，由于没有插入，则不被确认；

9. 重复 2-9 步，一直到所有的申请信息都被加入；
10. 点击 OK 按钮，确认并关闭窗口。

ⓘ提示：　电话、传真号码，邮政编码不会再次确认。

ⓘ PATENTIN 的新特点：　在 PATENTIN 3.1 中，点击 **OK** 后，编辑栏的数据如果和列表中的数据不同，则被加入列表中。

第五部分
序列数据

第五部分
序列数据

5.1 序列

用户在序列窗口（**Sequence Screen**，图 5-1）建立或修改序列，还可以建立或修改自定义密码子或生物体名称。还具有搜索功能，输入一段序列来搜索包含此序列所在文件的任务。当 PatentIn 启动时可以立即打开这个窗口。

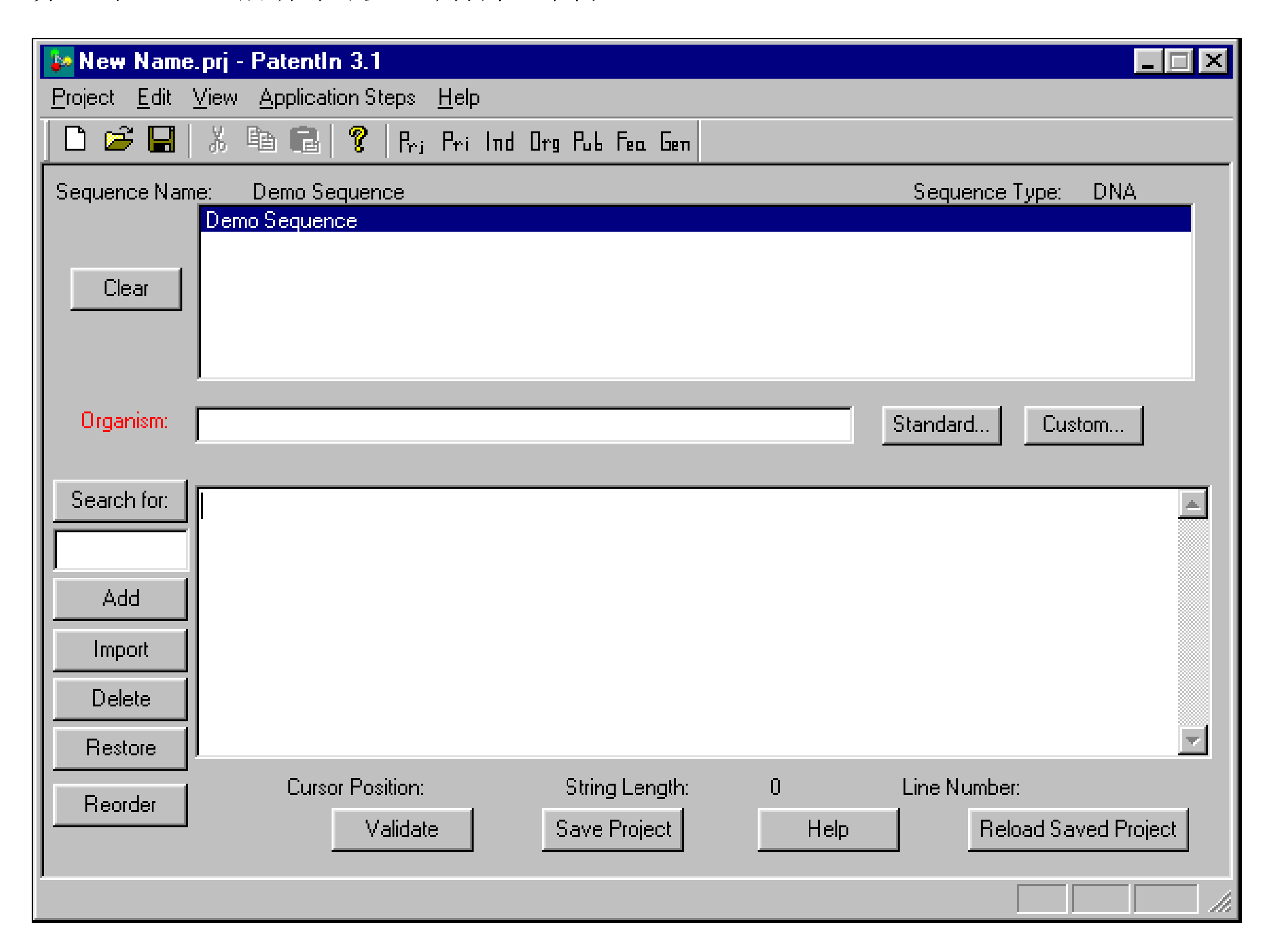


图 5-1: Sequence Screen

ⓘ注意：　　输入新的序列前，必须给它命名，参见 5.2 节。

选择序列进行编辑：

选中序列列表中一条序列。

可以看到以下的序列特征：

Cursor Pos 显示当前的光标位置，没有序列时这个区域是空白的。

String Length -显示行中序列的长度。

Line Number -表示光标位置所在的行数。

5.1.1 选择标准生物体名称（**Organism**）

生物体名称选择窗口（**Selecting an Organism Screen**，图 5-2）中，用户可以从常见生物体列表中选择名称，程序具有匹配输入部分名称的功能以选择生物体的名称：

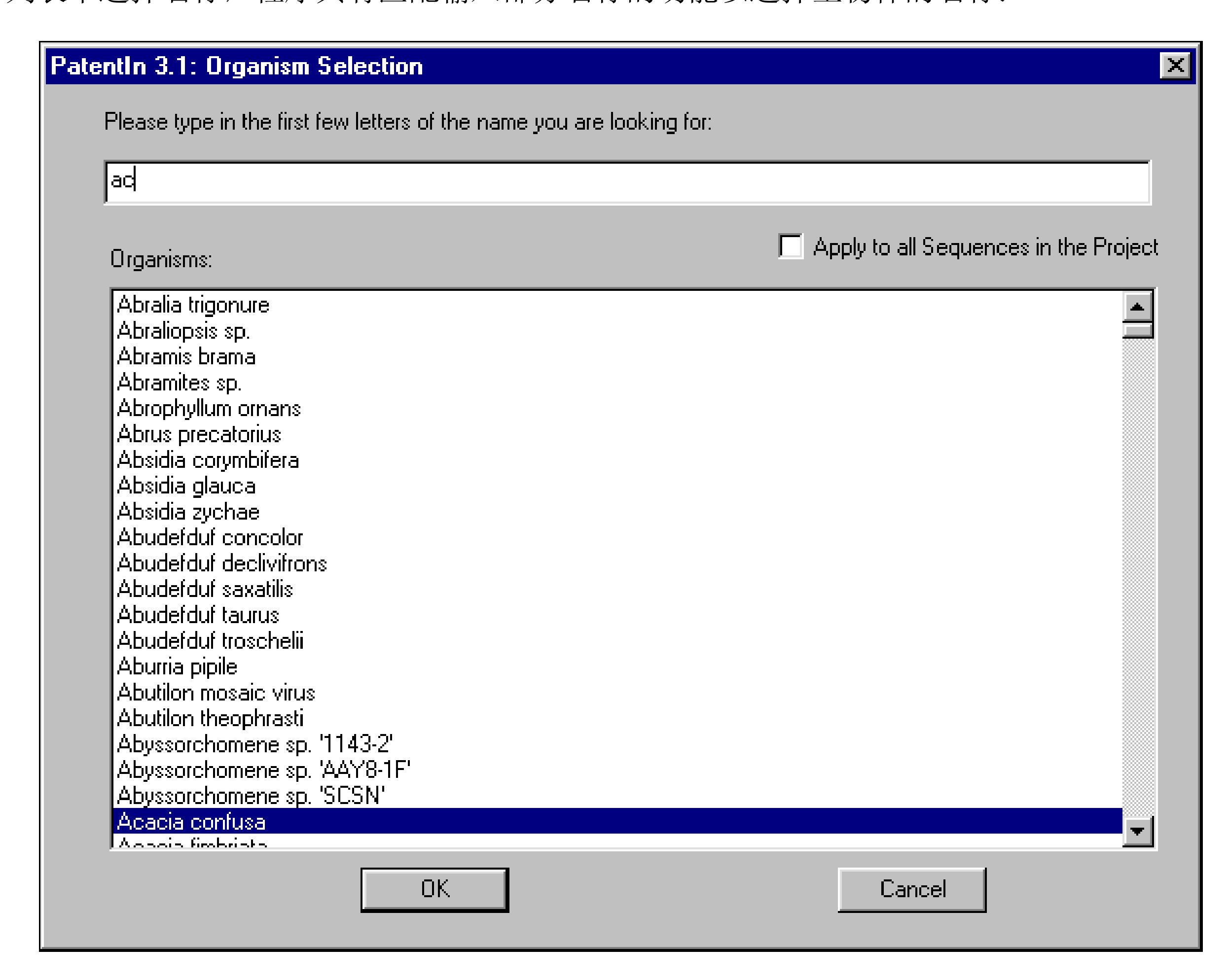


图 **5-2: Selecting an Organism Screen**

选择生物体的名称：

1. 点击 图 5-1中的 **Standard** 按钮；
2. 输入要寻找的生物体名称；

3. 点击 **Apply to all Sequences in the Project** 选框，可以赋予或撤消任务中的所有序列的生物体名称；
4. 点击 **OK** 按钮确认选中的生物体名称。

5.1.2 搜索序列

搜索特定的序列：

1. 在 **Search for** 按钮下方的编辑区输入特定的字符串（比如某特征）（图 5-1）；
2. 点击 **Search** 按钮，光标则从当前停留位置移动到第一个包含字符串的位置处；

提示:　　每次搜索的长度被限制或截短至 60 个字符。

5.1.3 清除屏幕内容

1. 点击 **Clear** 按钮（图 5-1）即可清除屏幕上关于某条选中序列的信息；

5.2 添加序列

点击序列窗口（**Sequence Screen**，图 5-1）中的 **Add** 按钮可以输入序列的名称，并从单选的按钮中选择序列类型（图 5-3）。

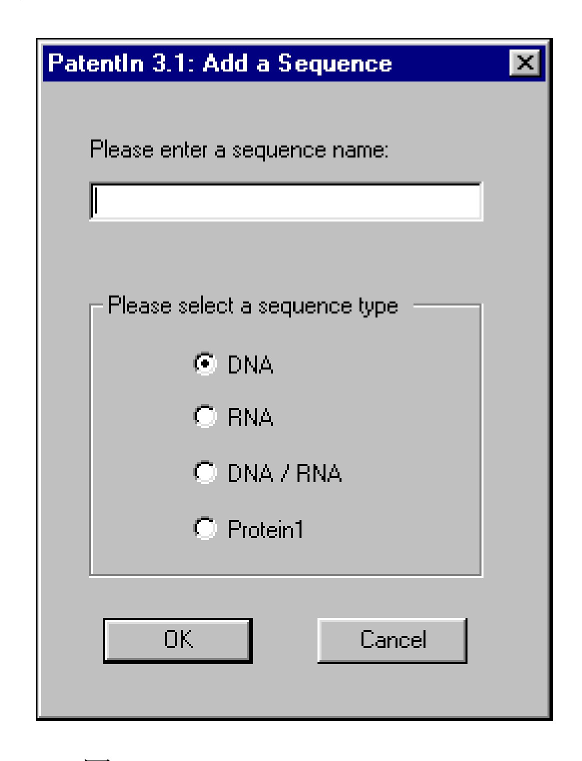


图 **5-3: Add A Sequence Screen**

添加序列的方法：

1. 点击序列窗口中 **Add** 按钮，出现添加序列窗口（**Add A Sequence Screen**，图 5-3）；
2. 在对话框中输入序列名称；
3. 点击需要的序列类型前的单选框◉；
4. 点击 **OK**。

在窗口底部的编辑区可以输入序列，如果使用的操作系统是 Windows 95，则最多只能输入 64，000 个字母。而 WindowsNT，98，2000 则上限超过一百万个字符。

如不使用序列编辑器，而使用导入文件功能可以突破这些限制来创建并编辑序列。

5.3 导入序列

序列窗口（图 5-1）中的 **Import** 按钮，可以从一个文件或任务中输入序列，甚至通过选择三个单选按钮来导入 PatentIn 2.1 版的任务。

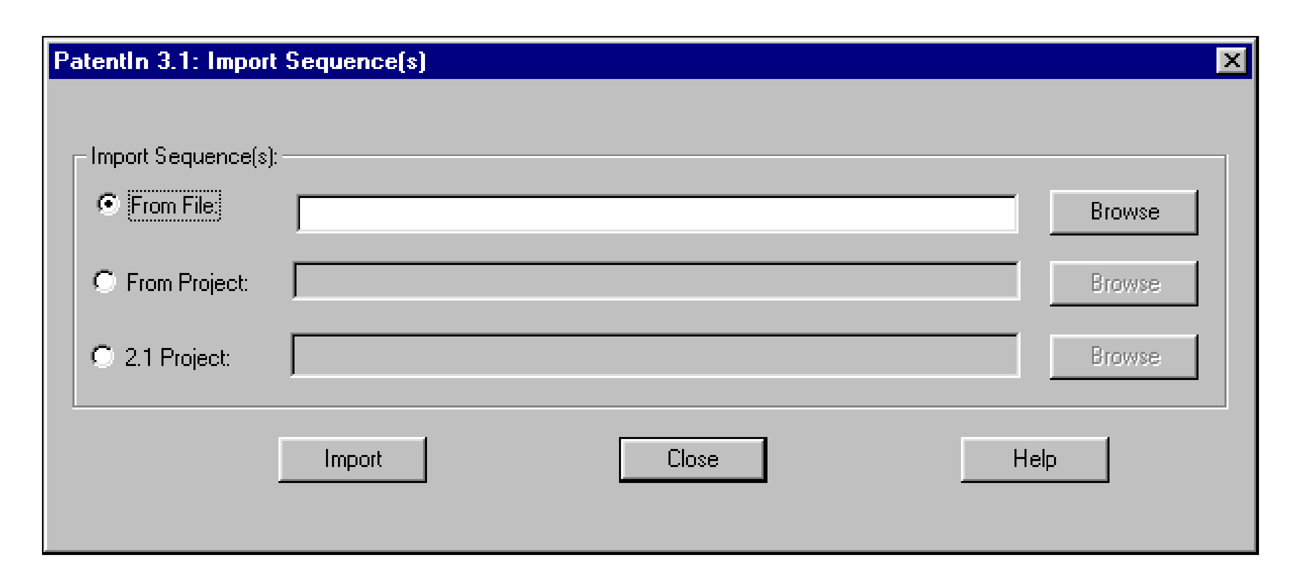


图 5-4: Import Sequence(s) Screen

1. 从文本文件中输入序列，点击 **Import** 按钮，然后点击 **From File** 单选按纽（图 5-4）；
2. 从其他任务中输入序列，点击 **Import** 按钮，然后点击 **From Project** 单选按纽（图 5-4）；
3. 从 PatentIn 2.1 的任务中输入序列，点击 **Import** 按钮，然后点击 **2.1 Project** 单选按纽（图 5-4）；

4. 每个单选框有 **Browse** 按钮帮助用户查找文件夹和文件名称，并处理选定的不同文件的格式。

ⓘ 提示: 输入 **2.1** 任务需要安装其他软件，见 5.3.4 节。

ⓘ 提示: 选择了 Protein/3 项，只能从仅由氨基酸缩写名称组成的文件中输入数据，参见附录 D：核酸三联体密码和单或三字母氨基酸符号对应表，PRT/3 栏。PRT/3 字符转换为 PRT/1 字符便于在序列编辑器中使用。

5.3.1 导入 PatentIn3.x 的多序列数据的格式

序列文件是一个包含一或多条序列的 ASCII 文本文件，每个多序列数据文件必须以一个符合 1.5.3.1.1 中描述的格式的标题开头，并且标题必须是行中第一条非空的文字。

5.3.1.1 序列标题

整个标题必须在单行中（表 5-2）。

<*SequenceName;SequenceType;OrganismName*>

表 5-2: Header for Sequence

SequenceName	序列名称
SequenceType	序列类型，下列类型之一: • DNA • RNA • DNA/RNA • Protein/1 • PRT • PRT/1 • PRT1 • Protein/3 • PRT/3 • PRT3
OrganismName	生物体名称，是可选的。如果省略了，标题则象：<*SequenceName;SequenceType*;>。

ⓘ 提示: 注意：可以看到一些分号，这是必须要的。

5.3.1.2 序列数据

序列数据紧跟在标题行的下面，内容是与序列类型相符的一串字母，可以占多行但不允许包含空格，空格表示序列数据的末尾。

序列数据以一或若干个空格截止，或者以另外一个标题开头，序列的末尾或开头可以有若干空行。

比如象下面的 2 条序列组成的 **ASCII** 文本格式的样品序列数据（**Sample ASCII Sequence Data**，图 5-5）：

```
< First Sequence;DNA;Abies alba>
ttttcttattgtttctcctactgcttatcataatgattgtcgtagtggcttcctcatcgt
ctcccccaccgcctaccacaacgactgccgcagcggattactaatagtatcaccaacagc
ataacaaaaagaatgacgaagagggttgctgatggtgtcgccgacggcgtagcagaagga
gtggcggagggg

< Second Sequence;RNA; >
uuuucuuauuguuucuccuacugcuuaucauaaugauugucguaguggcuuccucaucgu
cucccccaccgccuaccacaacgacugccgcagcggauuacuaauaguaucaccaacagc
auaacaaaaagaaugacgaagaggguugcugauggugucgccgacggcguagcagaagga
guggcggagggg
```

图 5-5: Sample ASCII Sequence Data

5.3.2 PatentIn3.1 输入的单一序列数据格式

序列文件是一个包含一条或多条序列的 ASCII 文件，单一序列不必一定要标题，如果没有标题，程序会询问序列类型（图 5-6），而且文件被认定为单一序列文件。

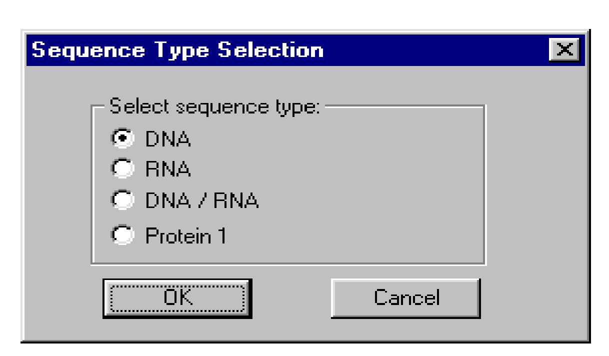


图 5-6: Sequence Type Selection Screen

ⓘ提示：　PatentIn2.1 中 ASCII 序列文件使用".gbs"作为文件扩展名。

使用导入功能时，一个小窗口显示已经分析过的序列的条数。下图中（图 5-7）表明了一个名称为 SEQ1000 的多序列文件目前已经包含了 513 条序列。

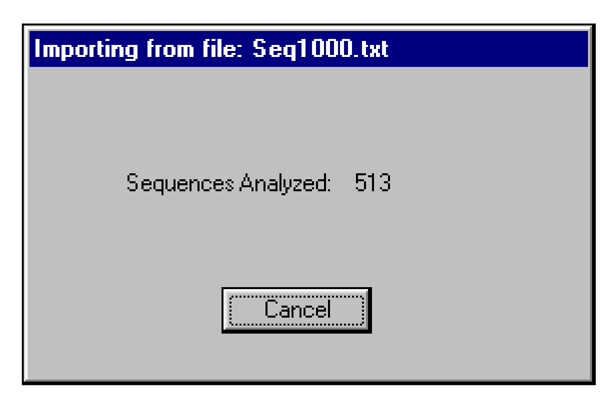


图 5-7: **Sequences Being Imported Screen**

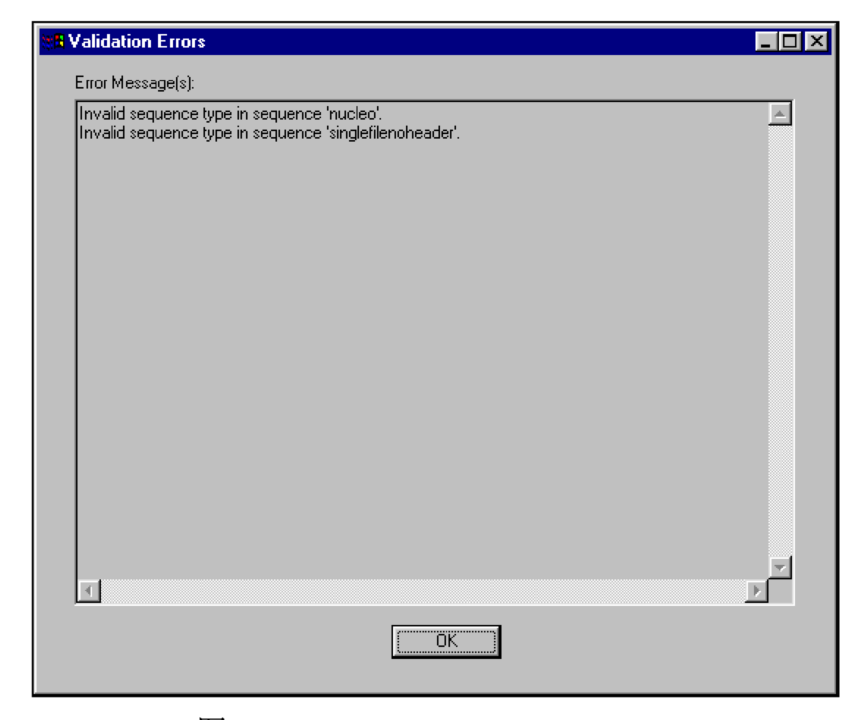


图 5-8: **Validation Errors Screen**

如果确认时发生错误会出现一个确认错误窗口，如下面的例子（图 5-8）中，singlefilenoheader 这个文件被作为 DNA 文件而不是蛋白质文件，Nucloe 只是名字象核酸，并非指序列。

5.3.2.1 从任务中导入序列

PatentIn 中，用户可以从 PatentIn3.x 任务中导入序列，见序列导入窗口（**Import Sequence(s) Screen**，图 5-4）。

1. 在序列窗口（图 5-1）中点击 Import 按钮，然后点击 **From Project** 可以从任务中导入序列；

2. 利用 **Browse** 按钮，可以查找文件夹和文件；
3. 选择一个任务时，任务所包含的序列就会显示出来；
4. 🖰 点击序列就可以导入；

ⓘ提示： 按住 Ctr 键不放，可以用鼠标点选多条序列。

5. 或者🖰点击 **Select All** 按钮，选择所有序列（图 5-9）。

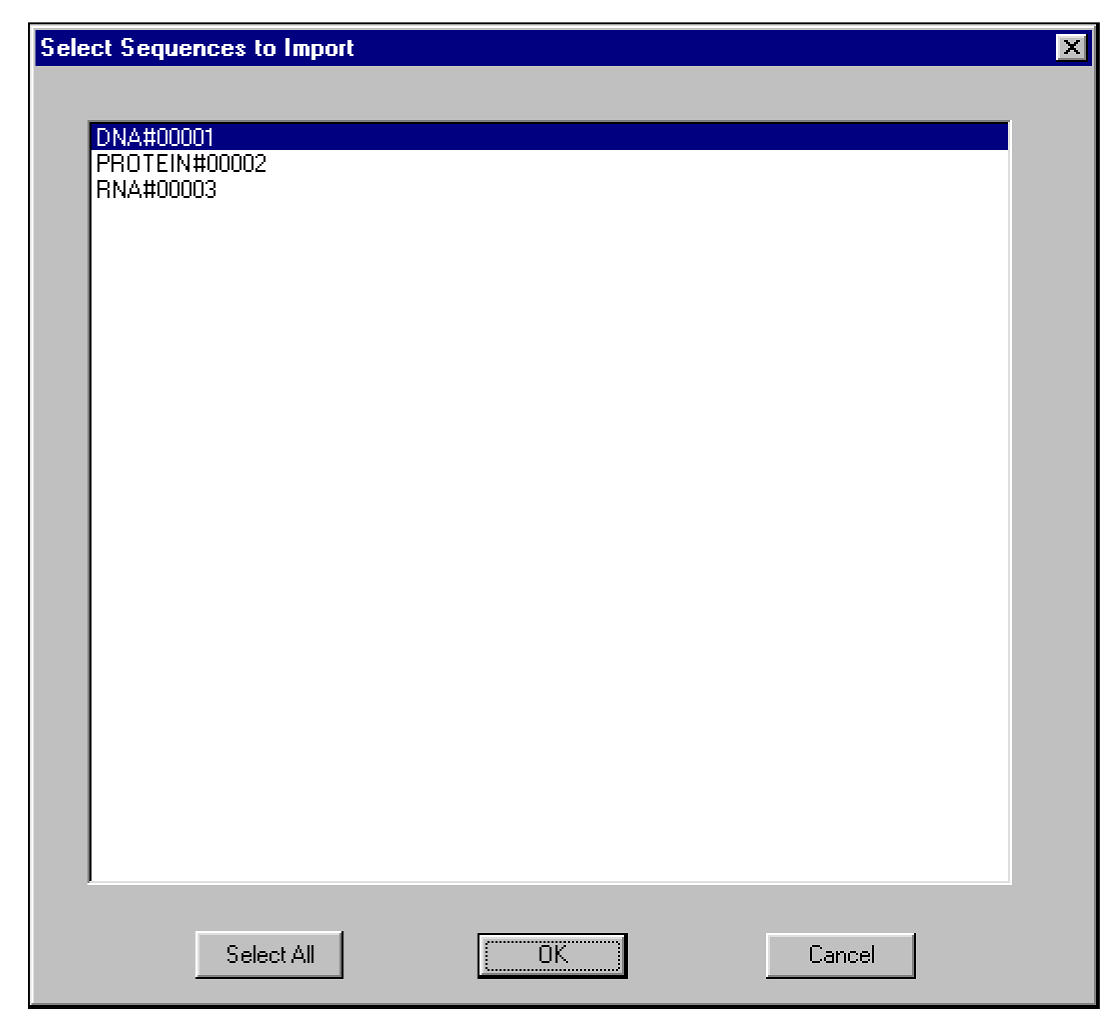


图 **5-9: Select Sequences From Project Screen**

5.3.2.2 导入 PatentIn 2.1 版的任务

1. 🖰 点击序列窗口中的 **Import** 按钮，再🖰选中 **2.1 Project** 单选框，就可以从 PatentIn2.1 中导入序列；
2. 利用 Browse 按钮，可以查找文件夹和文件。要注意的是，这里选中的是.dbf 格式的文件，而不是任务或文件，见（图 5-10）。

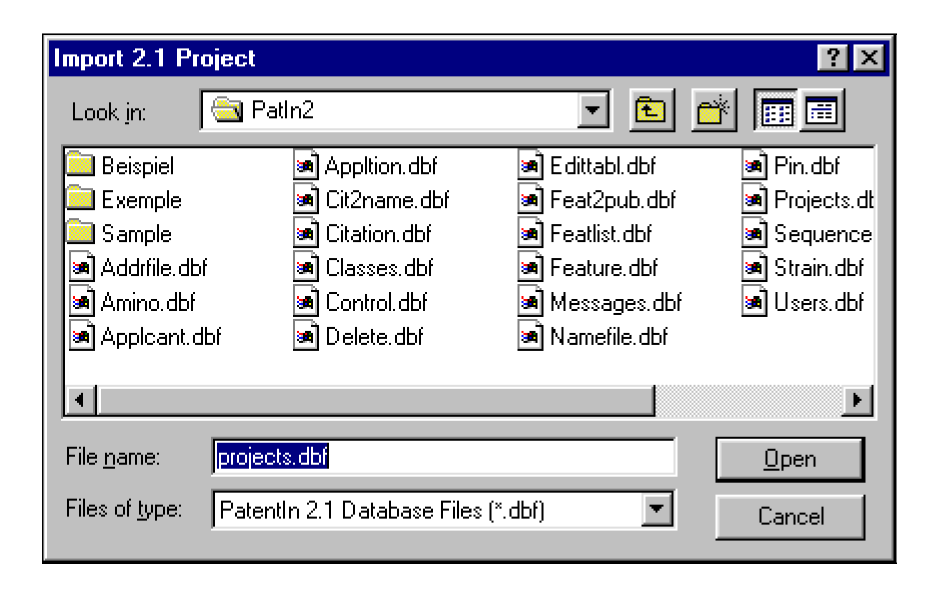


图 5-10: Browse Window for 2.1 Project Import

3. 选择 PatentIn2.1 的任务后，会显示全部任务的列表，见（图 5-11）。

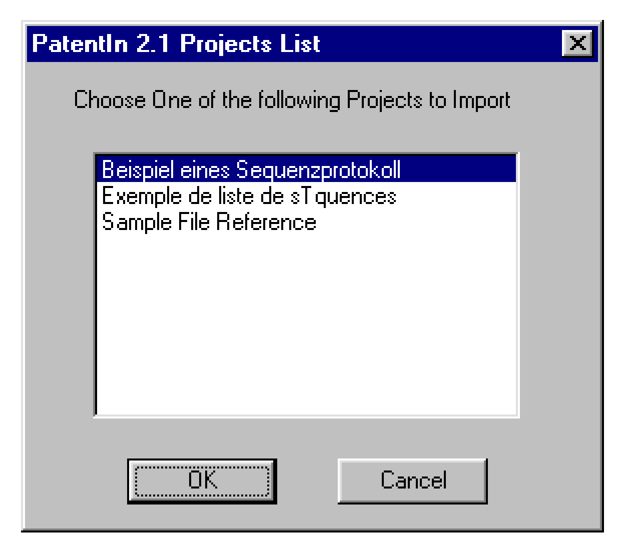


图 5-11: PatentIn 2.1 Projects List

4. 🖱 点击要导入的任务；
5. 🖱 点击 **OK** 按钮。

5.4 拷贝序列

PatentIn 具有典型的 Windows 类型的编辑特色。

拷贝序列：

1. 🖱 选中要拷贝的序列；
2. 🖱点击 Edit 按钮，然后🖱 点击拷贝（**Copy**， 图 5-12）；

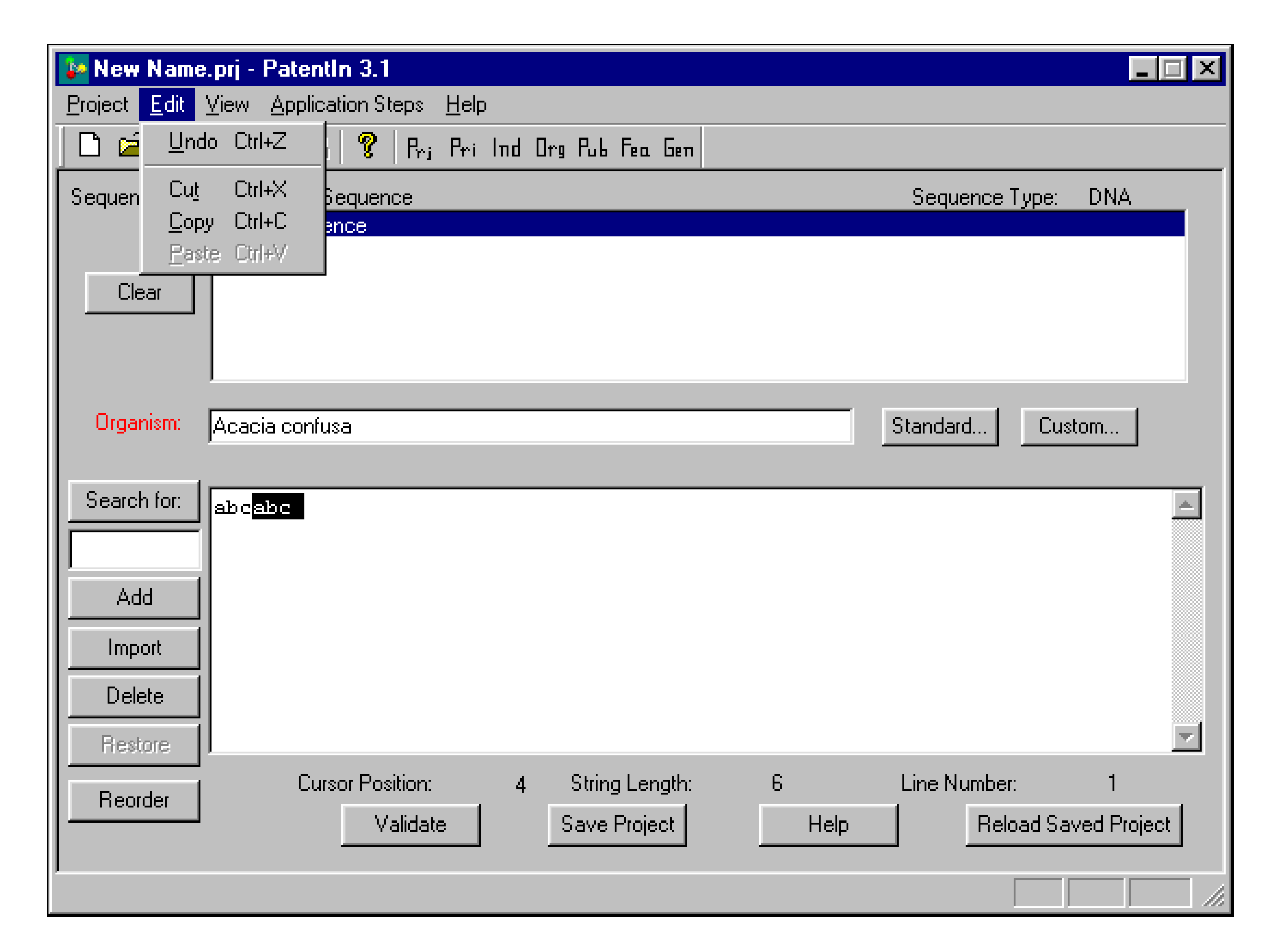


图 **5-12: Edit Menu**

5.5 粘贴序列

粘贴序列:

1. 将光标放到待粘贴的位置;
2. 🖱点击 **Edit**，然后🖱点击粘贴（**Paste**）。

5.6 删除序列

方法:

1. 将光标放到待删除的位置;
2. 🖱点击删除（**Delete**）按钮。

5.7 取消删除

当序列被删除后，只要目前的任务没有关闭，还可以恢复序列。

方法：

1. 点击 **Restore** 按钮，出现序列恢复窗口（**Sequence Recovery Screen**，图 5-13）；
2. 选择要恢复的序列；
3. 点击恢复（**Restore**）按钮。

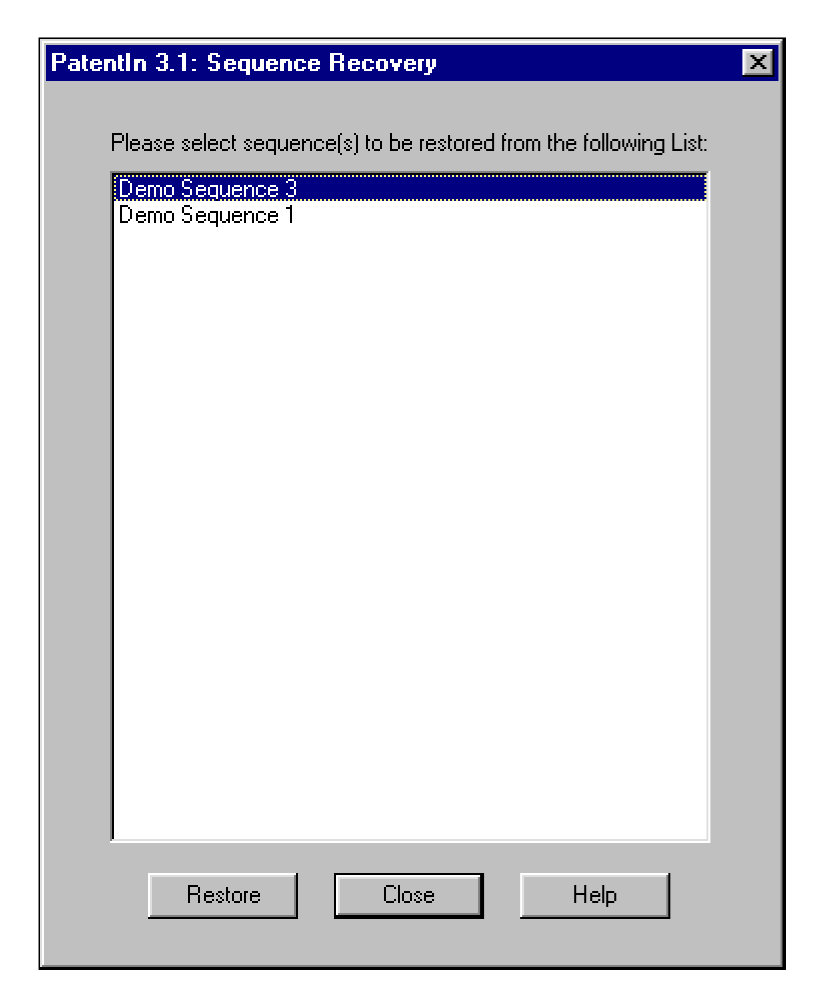


图 **5-13: Sequence Recovery Screen**

5.8 序列重新排序

重新排序窗口（**Reorder Sequences Screen**，图 5-14）中可以比较当前序列的次序以及变动后序列的次序。窗口左边是当前序列，按输入到申请中的次序排列；右边是变动后的次序，是从左边选择一组连续排列的序列，放到右边某条序列的后面。

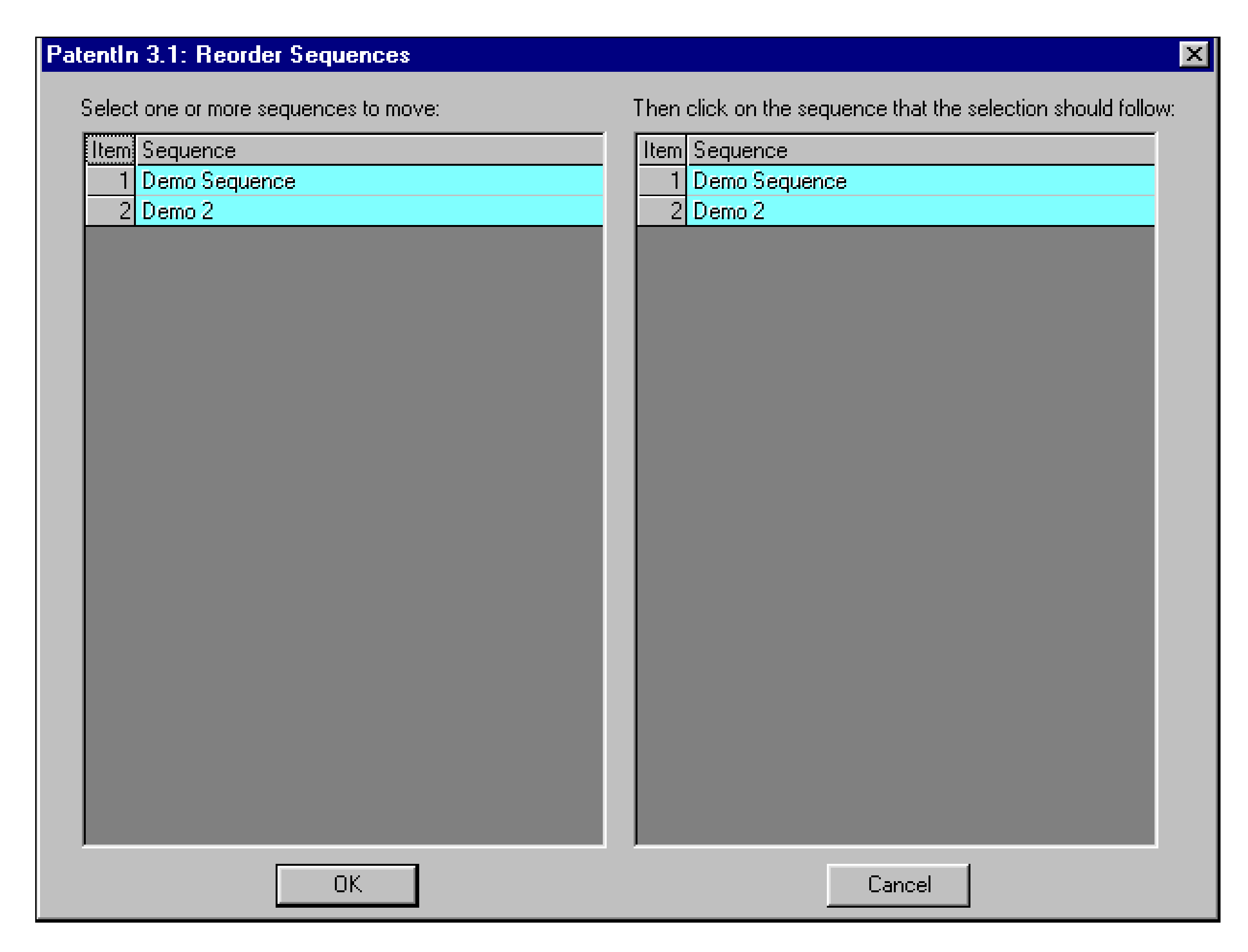


图 **5-14: Reorder Sequences Screen**

方法：

1. 选中左边序列的一或几行；
2. 点击插入位置前的序列；
3. 重复 1-2 直到序列按照要求的次序排列。

5.9 确认序列

方法：

1. 在序列窗口（**Sequence Screen**，图 5-1），点击 **Validate** 按钮，会有消息窗口提示出错，否则状态栏会出现确认的 **OK;**

提示：　　名称被选中的序列的数据得到确认。

5.10 保存序列

在序列窗口（**Sequence Screen**，图 5-1），点击**保存任务**（**Save Project**）按钮，工作将按照目前的状况保存。

提示:　　一定要记住，当处理非常大或非常复杂的任务时，及时保存任务可以避免由于系统崩溃导致几小时的工作功亏一篑。

5.11 打开保存过的任务

使用 **Reload Saved Project** 按钮可以迅速地加载保存过的任务（图 5-1）。

5.12 添加自定义的密码

通过**自定义密码输入窗口**（**Custom Codons Input Screen**，图 5-15）可以向标准密码列表中添加用户自定义的密码。从 **Application Step** 菜单（图 4-1）中选择**自定义密码**（**Define Custom Codon**）项可以打开下面的窗口。

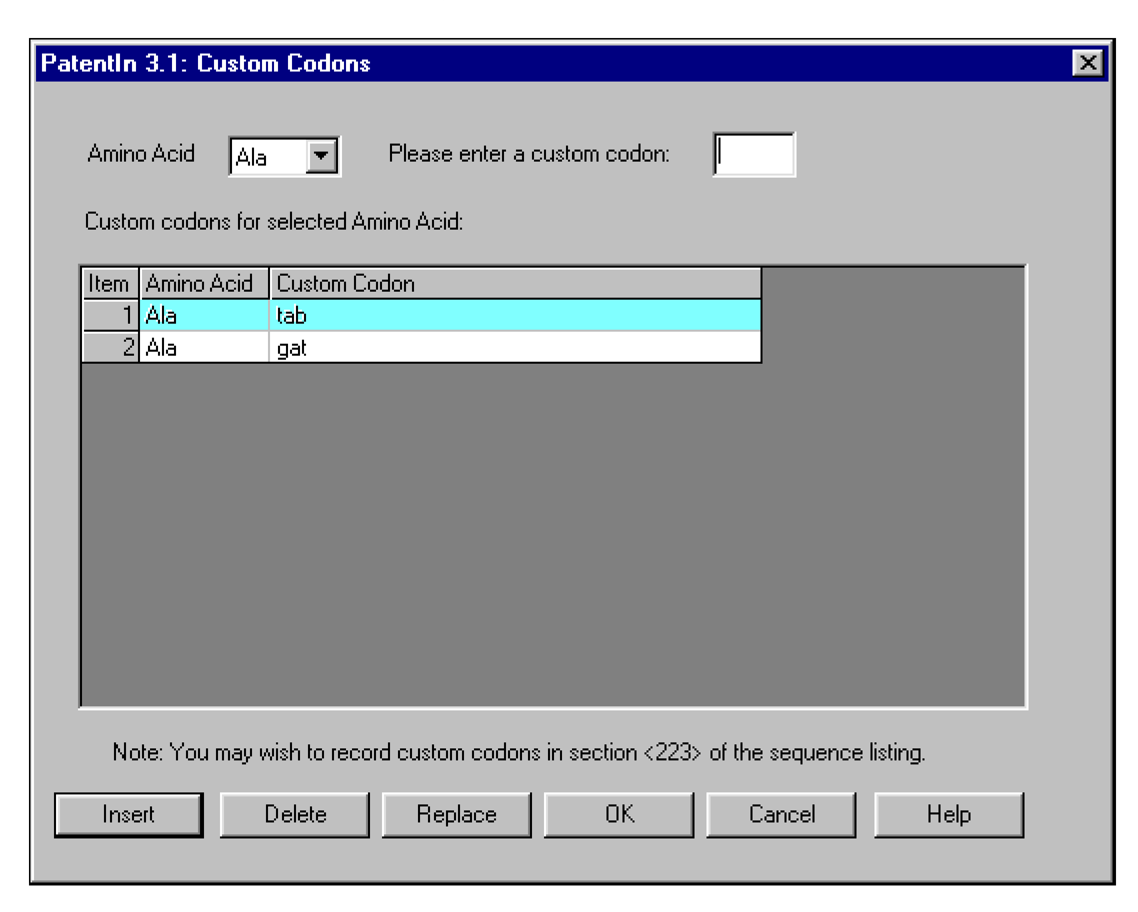


图 5-15：Custom Codons Input Screen

添加自定义密码：

1. 选择 **Application Step** 菜单（图 4-1）中的 **Define Custom Codon** 项；
2. 从下拉菜单列表中 选择 **Amino Acid**（图 5-16）；
3. 输入自定义密码（**Custom Codon**）；
4. 点击 **Insert** 按钮。

ⓘ 提示: 目前窗口的形式已经发生改变，允许用户在一个窗口中看到所有自定义的密码，不用一个个地选择每个氨基酸。

删除自定义的密码：

1. 点中列表中的自定义密码（**Custom Codon**）；
2. 点击 **Delete** 按钮。

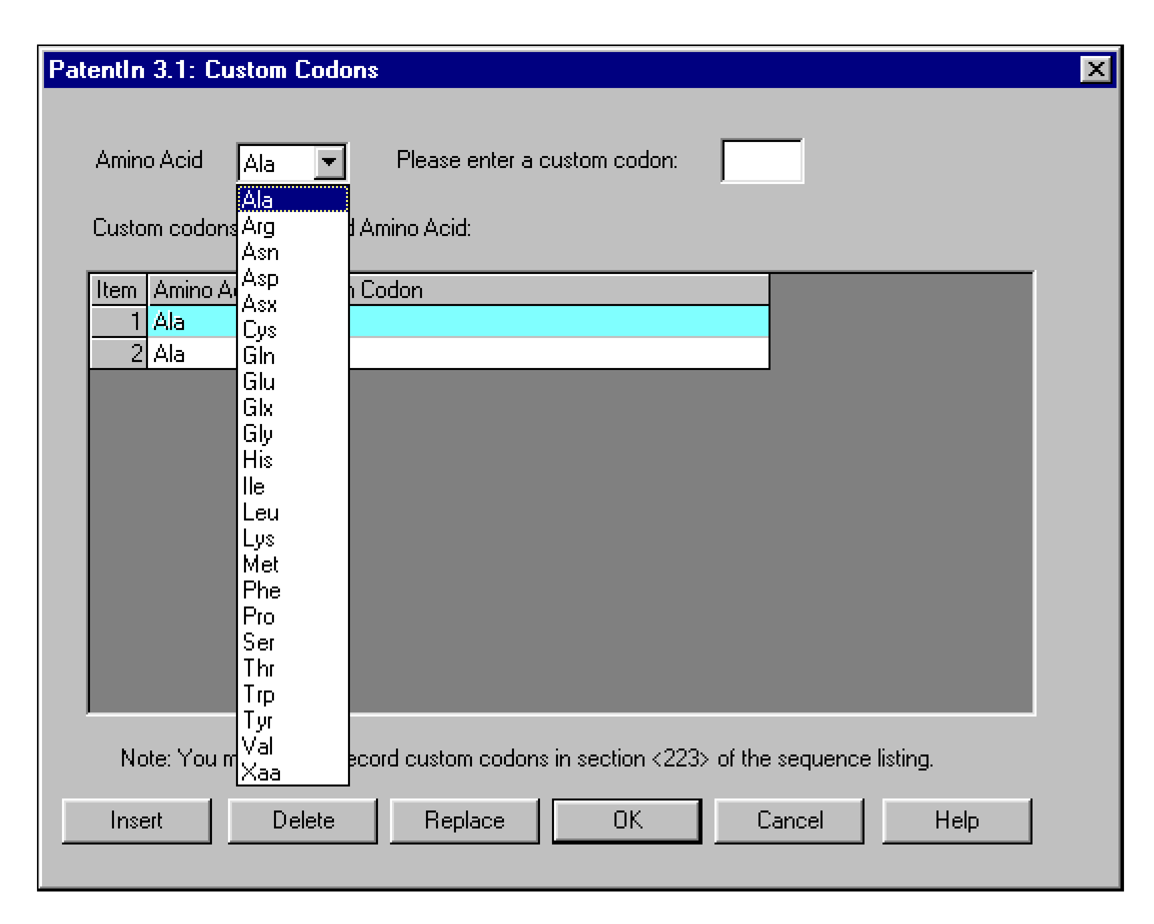


图 5-16: Amino Acid Drop Down List Screen

5.13 添加自定义生物体名称

在自定义生物体名称输入窗口（**Custom Organism Input Screen**，图 5-17）中，可以向列表中添加用户自定义的生物体名称，还可以选择自定义生物体名称输入到**序列窗口**（图 5-1）中。方法是从序列窗口中选择 **Custom** 按钮，用户添加生物体名称，然后向列表中添加或删除生物体名称。

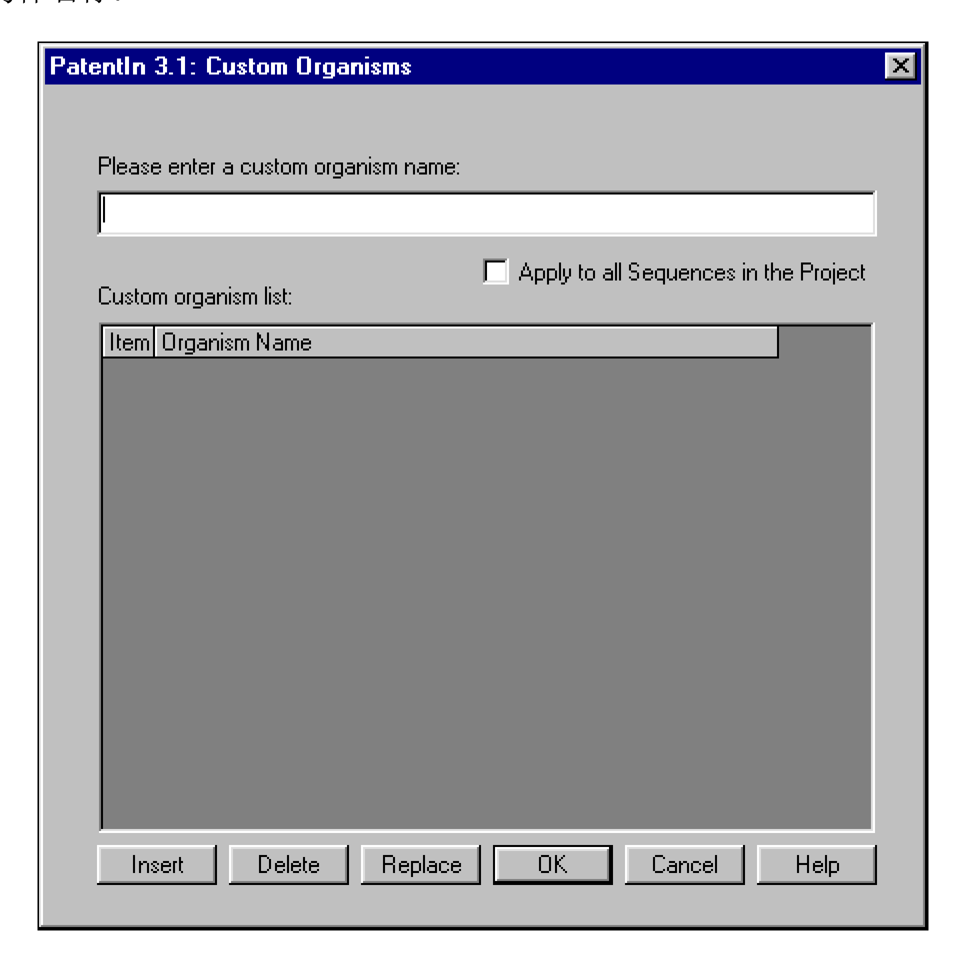


图 5-17: Custom Organism Input Screen

添加自定义生物体名称方法：

1. 选择序列窗口中自定义（**Custom**）按钮；
2. 输入生物体的名称；
3. 点击 Insert 按钮。

删除自定义的名称：

1. 点中列表中的自定义名称；
2. 点击 **Delete** 按钮。

替换名称：

1. 点中列表中的自定义的名称；
2. 输入自定义的生物体名称；
3. 点击替换（**Replace**）按钮。

从序列窗口输入自定义的名称：

1. 选择生物体，出现输入自定义名称框（“Please enter a custom organism name:”）；
2. 点击 **OK**。

给任务中所有的序列命名：

1. 选中 **Apply to all Sequences in the Project** 选框；
2. 点击 **OK** 按钮。

5.14 人造序列或未知名称的生物体

对于人造序列或未知名称的生物体，必须附加关于生物体的说明。在 **Organism Name** 栏填入 Aritfical Sequence 或者 Unknown 后，再到自动弹出的框中填写。

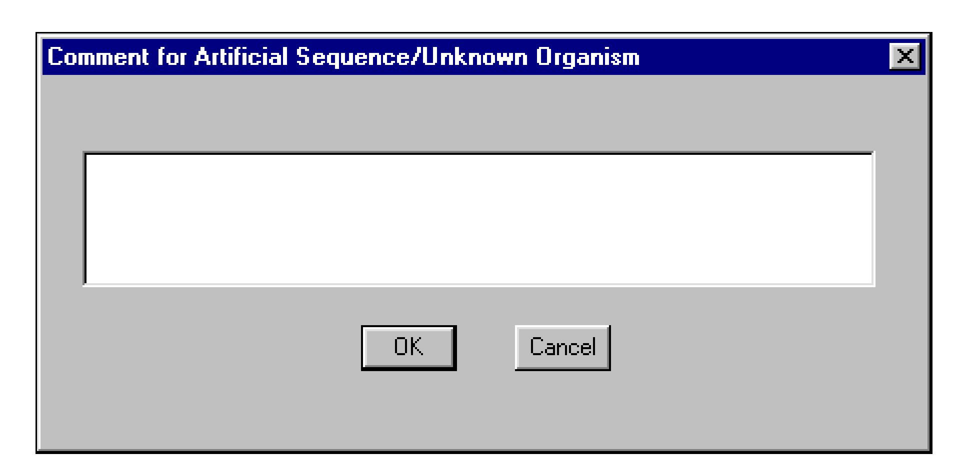


图 **5-18: Artificial Sequence/Unknown Organism Comment**

1. 选择序列窗口中的 **Standard** 按钮（图 5-1）；.
2. 从标准生物体名称列表中选择 Aritfical 或者 Unknown；
3. 按 Tab 键移动到另外一格，说明框中（图 5-18）自动弹出，要求填入生物体的定义；
4. 输入关于生物体正确的说明，当创建序列列表时，这部分的信息出现在**<223>**栏；

提示:　　也可以通过选择 Application Steps | Artificial Sequence/Unknown Comment（图 4-1）这些步骤实现。

提示:　　这部分应该是描述性的而且尽量简洁。

ⓘ 提示:　　序列列表中的<223>栏可以用 Artificial Sequence/Unknown Organism 或者 misc_feature 来代替，是相关的说明或其他信息。

第六部分
特征数据

第六部分
特征数据

6.1 序列特征

序列特征窗口（**Features Screen**，图 6-1）中可以建立或修改序列的特征。通过选择 **Application Steps Menu** 中的 **Feature Data**（图 4-1），或者选择 PatentIn 中的工具栏中的 **Fea** 实现将显示的特征赋予在序列窗口选中的序列。

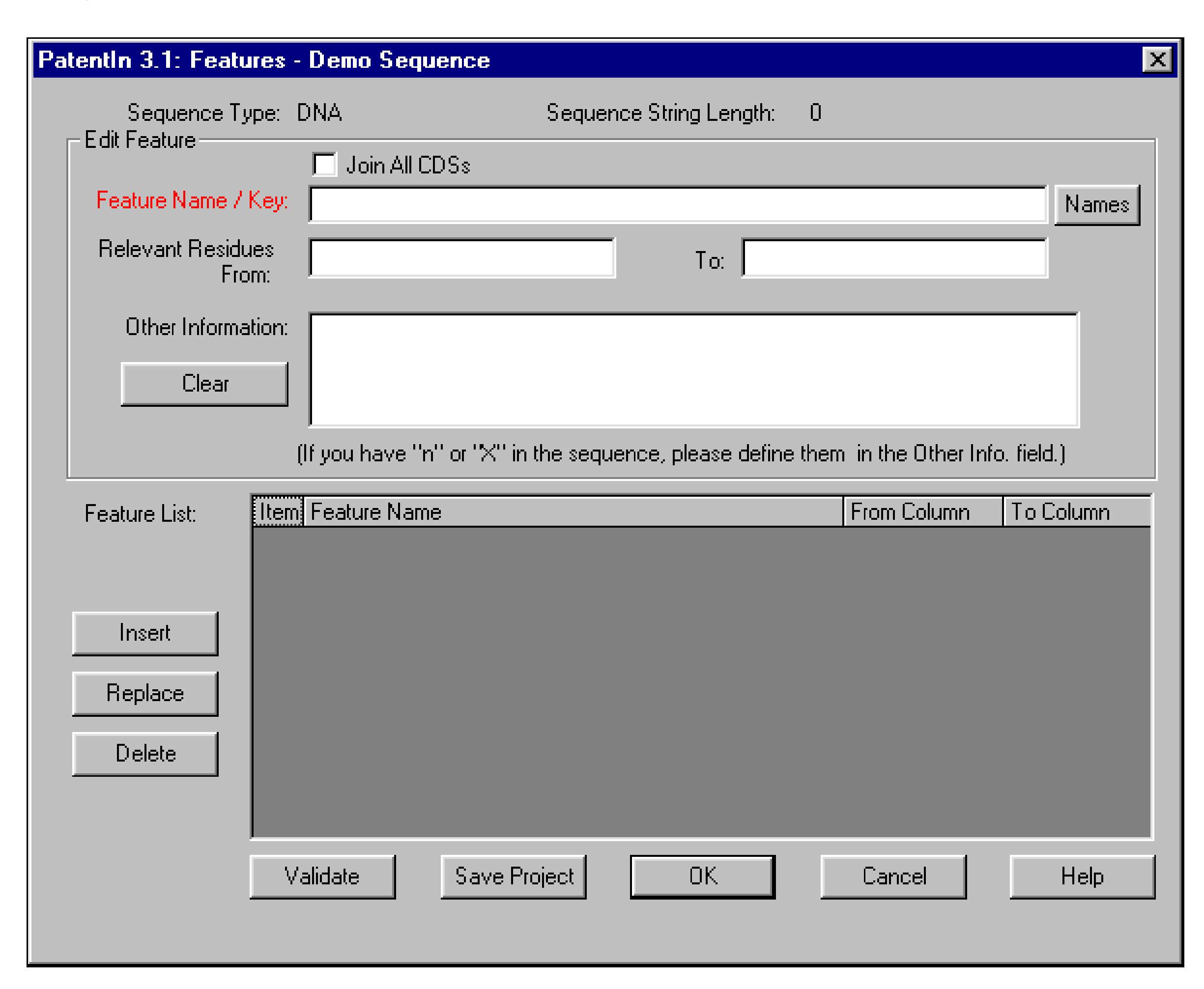


图 6-1: Features Screen

输入序列特征信息的方法：

1. 若序列含有多于一个 CDS，🖰则选中 **Join All CDSs**，以使所有 CDS 联系起来；
2. 🖰点击 **Names** 按钮获得 **Feature Name / Key** 的核酸序列名称的列表；
3. 🖮输入序列的起始（**Relevant Residue From**）、终止（**To**）位置的数字；

4. 点击 **Other Information** 框输入其他信息。这里在蛋白质序列中可以用 X，在核酸序列中可以用 n，Feature Name / Key 选择 misc_feature；
5. 点击 **Clear** 按钮可以清除窗口中的 **Edit Feature** 栏的信息；
6. 选中特征列表（**Feature List**）的一项，点击 **Insert** 按钮即加入一条序列特征；
7. 选中特征列表（**Feature List**）的一项，点击 **Replace** 按钮这条序列特征即被替换；
8. 选中特征列表（**Feature List**）的一项，点击 **Delete** 按钮即删除这条序列特征；
9. 点击 **Validate**，以确认输入的信息，这样输入或插入到列表中的数据得到确认。编辑栏是还没有确认的信息，是无效的；
10. 点击保存任务（**Save Project**），即可保存信息；
11. 点击 **OK**，确认并退出；
12. 点击 **Cancel** 按钮，消除输入的信息；
13. 点击 Help 按钮，可以获得帮助信息。

提示: **Join All CDSs** 选项存在与否，取决于特征列表中是否有多个 CDS。

提示: 序列列表中的<223>栏可以 Artificial Sequence/Unknown Organism 或者 misc_feature.来更新，是相关的说明或其他信息。

PATENTIN 的新功能: PatentIn 3.1 中，可以自动解决象 Xaa 之类的东西。

6.1.1 选择特征关键词

在特征名称/关键词选择窗口（**Feature Names / Key Selection Screen**，图 6-1）可以选择核酸序列类型的名称。

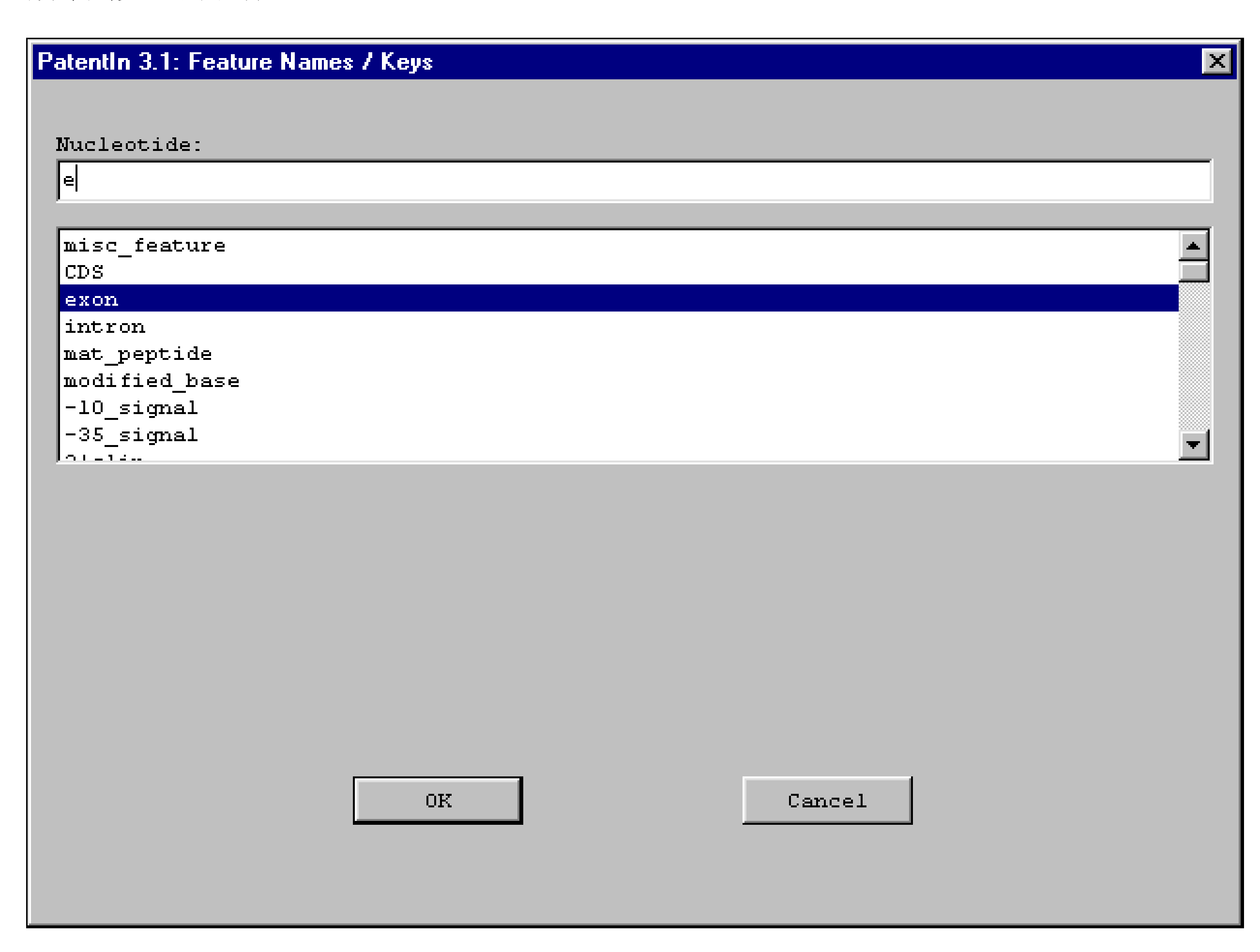


图 6-1: Feature Names / Key Selection Screen

选择核酸序列的名称：

1. 在特征窗口（**Feature Names / Key Selection Screen**，图 6-1），点选名称项；
2. 在 **Nucleotide** 栏输入或者
3. 点击向下箭头打开下拉列表（图 6-2），从列表中选择一个名称；
4. 点击 OK 确认选择，回到特征窗口（**Feature Screen**，图 6–1）。

6.1.2 修饰碱基的其他信息

选择 modified_base，特征名称/关键词选择窗口（**Feature Names / Key Selection Screen**，图 6-1）窗口自动扩大，出现关于修饰碱基的窗口。

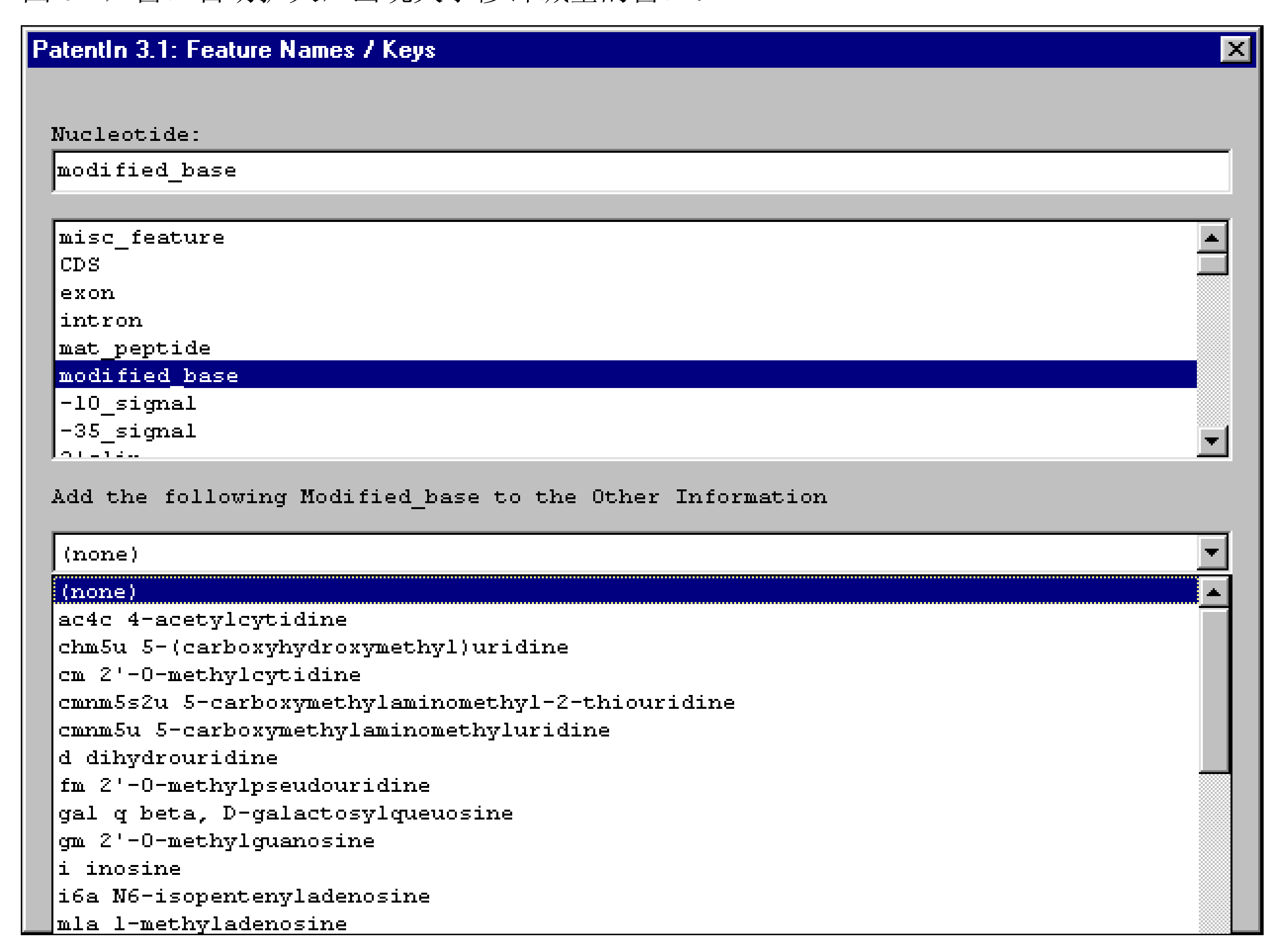


图 **6-1: Feature Names / Key Selection Screen with Modified Base**

添加关于修饰碱基的信息：

1. 选择 **Add the following modified_base to the Other Information** 框的下拉箭头；
2. 从列表中选择修饰碱基（modified_base， 图 6-1）；
3. 点击 OK（图 6-2）。

6.1.3 编码区的其他信息

当指定一条核酸序列的编码区时，DNA 序列以"混合"形式表示，DNA 链按密码子形式分开，下面对应着氨基酸。这就是选择“exon”特征的作用。但选择 CDS，使 PatentIn3.1 自动地产生多肽链，作为"额外的"序列。

6.1.4 进一步定义"n" 或 "Xaa"

如核酸、多肽序列中出现 n 或 Xaa，ST.25 规范要求有更多的定义。利用 Misc_Feature，在 Other Information 栏说明。PatentIn 把 n 的定义拷贝到"额外的"多肽链并翻译为 Xaa。

6.1.5 选择氨基酸名称

当选择 LIPID 时，特征名称/关键词选择窗口（**Feature Names / Key Selection Screen**，图 6-4）窗口自动扩大。

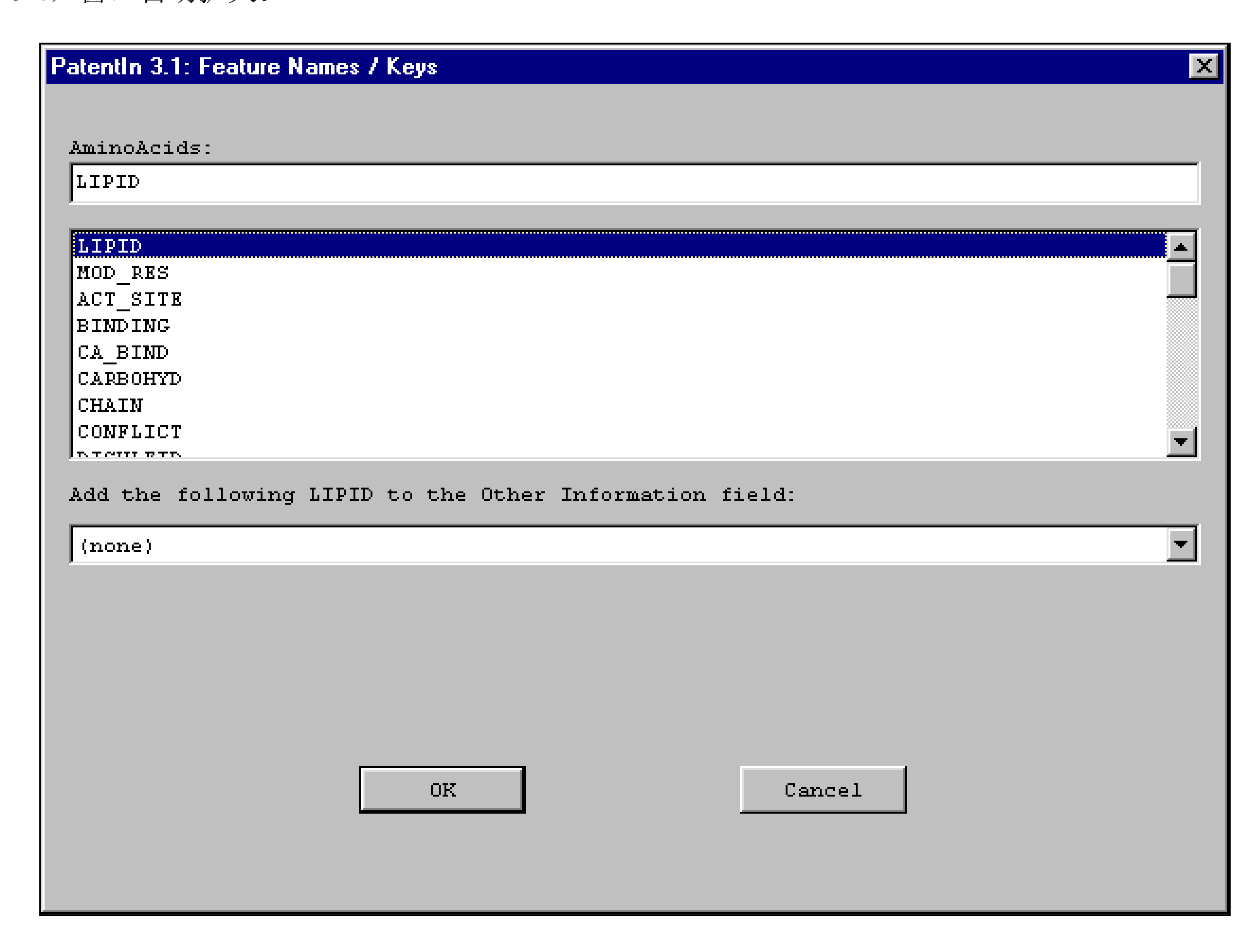


图 **6-4: Feature Names / Key Selection Screen**，已选中 LIPID

选择氨基酸名称：

1. 在特征窗口（Features Screen）中，🖰 选择 **Names** 按钮（图 6–1）；
2. 在 **Amino Acid** 栏输入，或者；
3. 🖰 点击向下箭头，打开下拉列表（图 6-5），从列表中选择一个名称；

4. 点击 OK 按钮，确定选择并回到特征窗口（**Feature Screen**，图 6-4）。

添加 **LIPID** 的信息：

1. 🖱 选择 **Add thefollowing LIPID to the Other Information** 框中的下拉箭头；
2. 🖱 选择列表中的 LIPID 信息；
3. 🖱 点击 **OK** 按钮（图 6-4）。

6.1.6 MOD_RES 的其他信息

选中 MOD_RES，特征名称/关键词选择窗口（**Feature Names / Key Selection Screen**，图 6-5）自动增加两行；

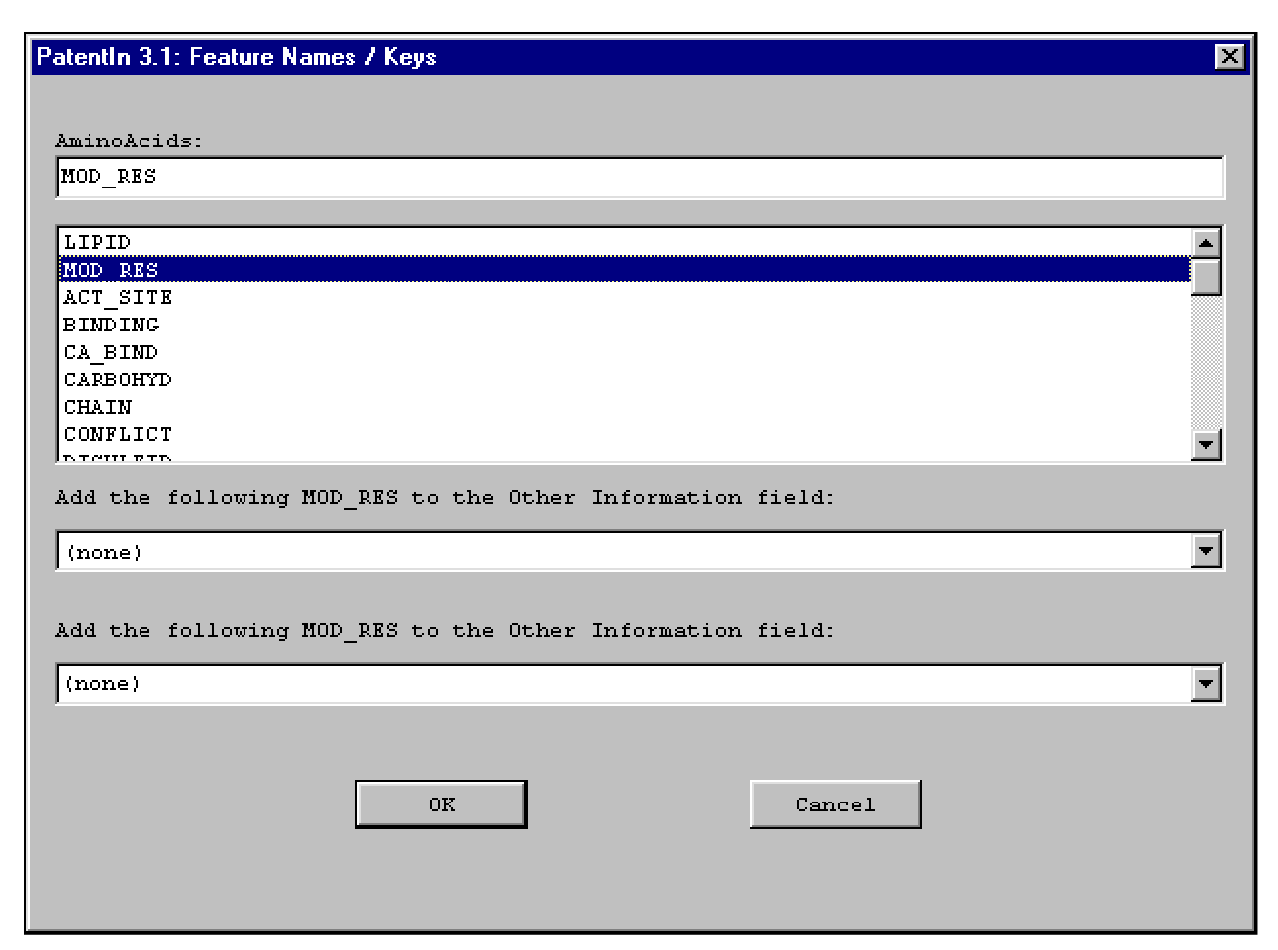


图 **6-5: Feature Names / Key Selection Screen**，已选中 MOD_RES

选择一个氨基酸名称：

1. 特征窗口（**Feature Screen**，图 6-1）中，🖱选择 **Names** 按钮；
2. 在 **Amino Acid** 栏开始输入，或者

3. 🖱 点击向下箭头，打开下拉列表（图 6-2），从列表中选择一个名称；
4. 🖱 点击 OK 按钮，确定选择并回到**特征窗口**（**Feature Screen**，图 6-4）。

增加 **MOD_RES** 的信息：

1. 🖱 选择 **Add the following MOD_RES to the Other Information field** 中第一个框中的下拉箭头；
2. 🖱 从列表中选择正确的信息（图 6-6）；
3. 🖱 点击 OK 按钮（图 6-5）；

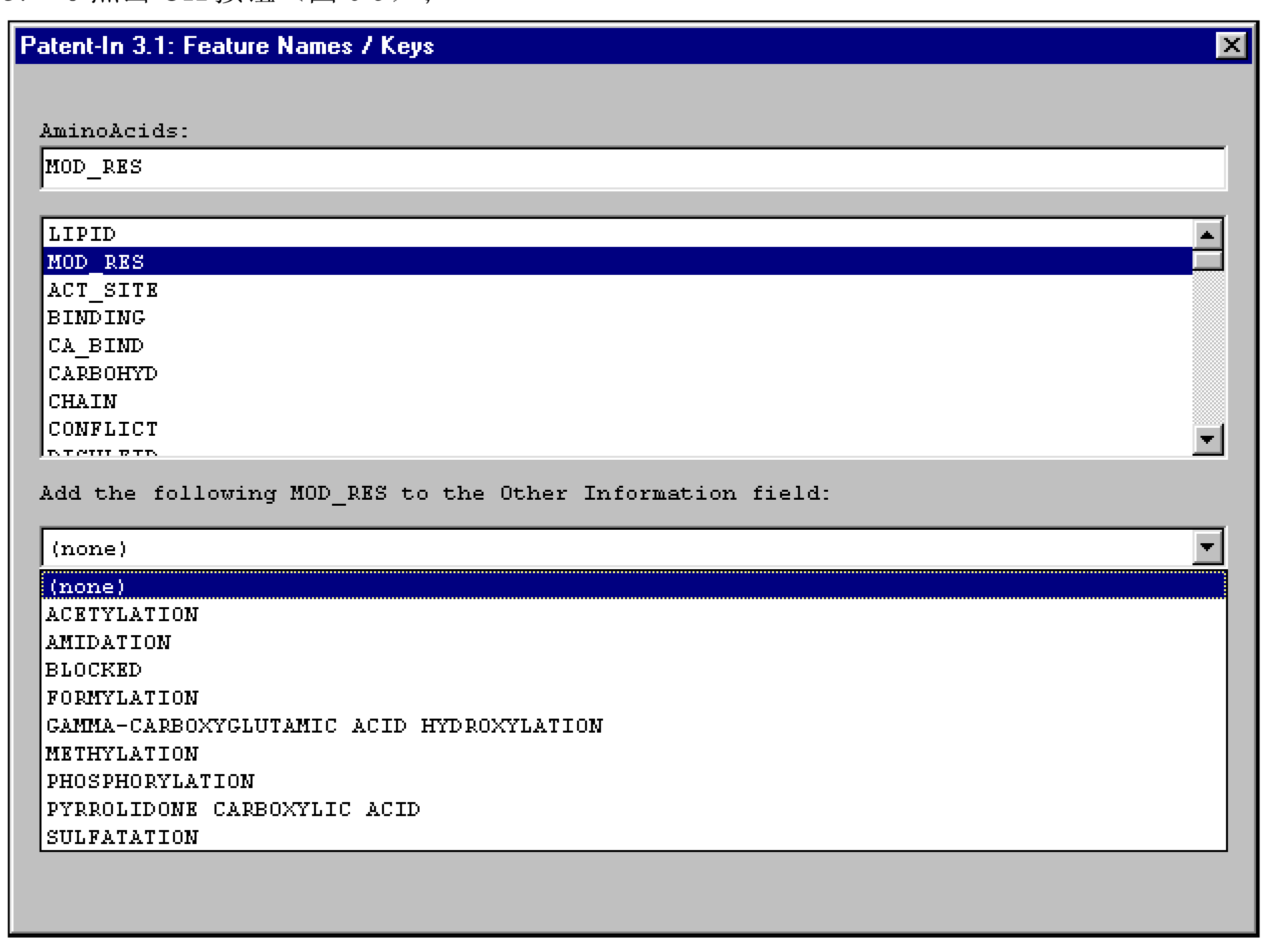


图 6-6: First MOD_RES Pull Down List

4. 🖱 选择 **Add the following MOD_RES to the Other Information field** 中第二个框中的下拉箭头；
5. 从列表中选择正确的信息（图 6-7）；
6. 🖱 点击 **OK** 按钮（图 6-5）。

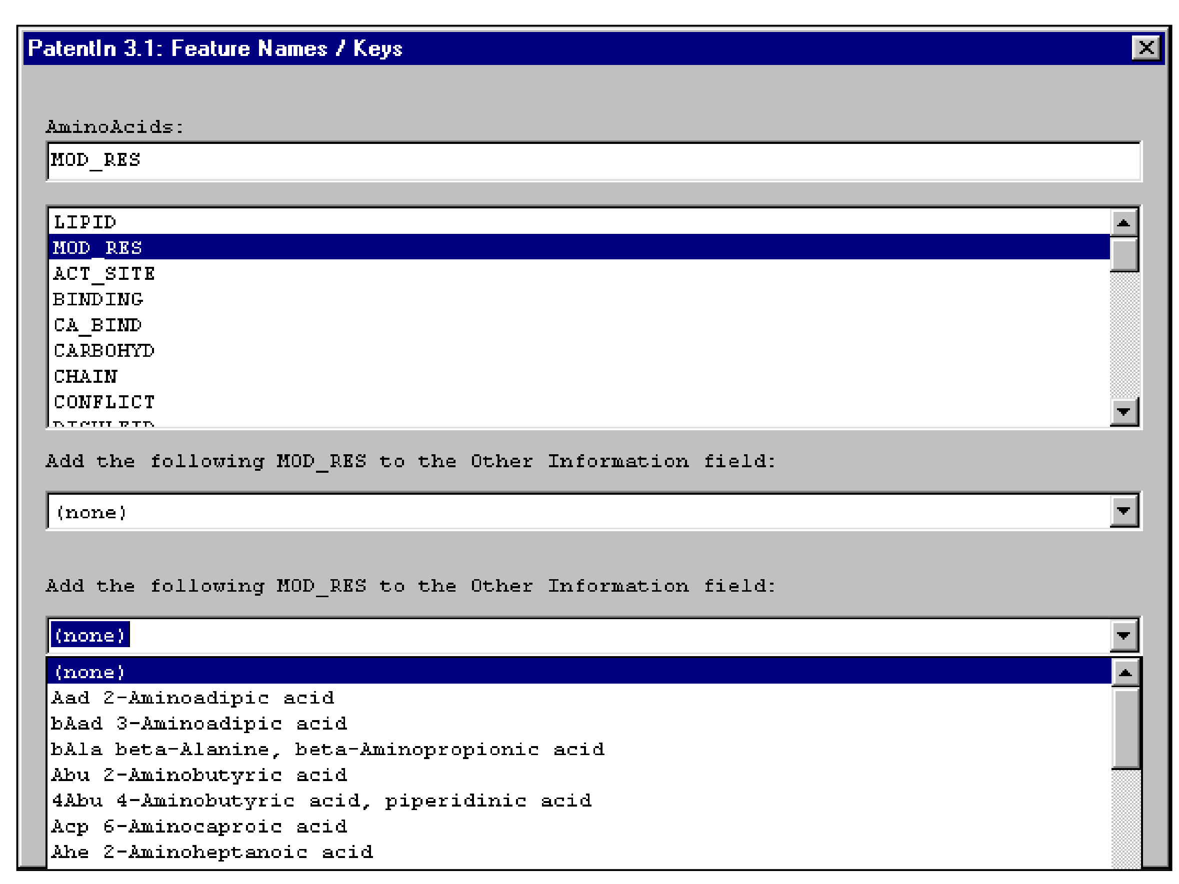


图 6-7: Second MOD_RES Pull Down List

第七部分
文献数据

第七部分
文献数据

7.1　文献类型窗口

公开文献类型界面（**Publication Type Screen**，图 7-1）提供了输入文献信息的 4 条途径：期刊、数据库、专利和论文。在 **Application Steps** 菜单中选择**文献数据**（Publication Data）进入。

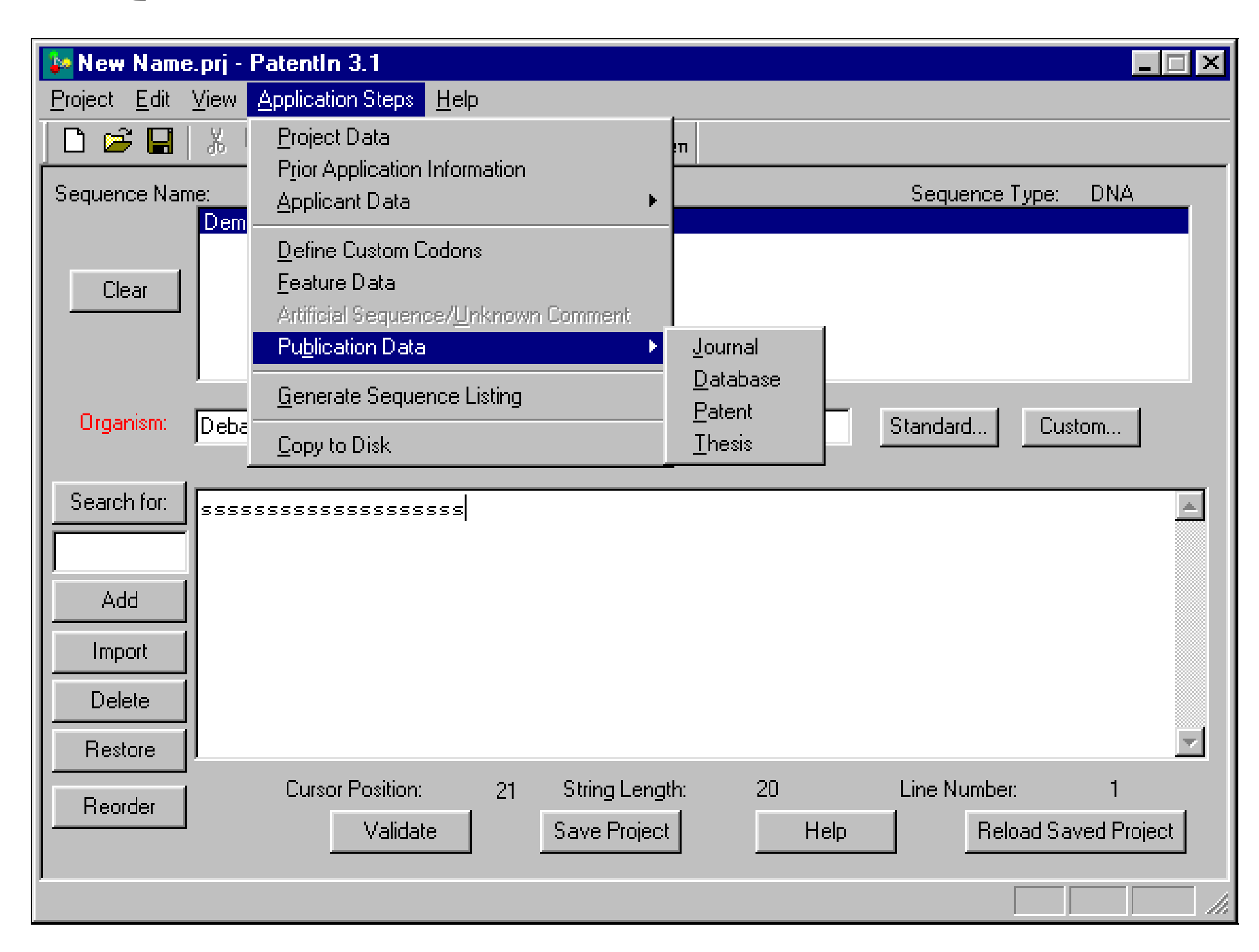


图 7-1: **Publication Type Screen**

选择文献类型：

1. 在申请程序（**Application Steps**）的下拉菜单中，🖰 点击**文献数据**（**Publication Data**）；

7.2 杂志文献信息

期刊文献信息窗口（**Journals Publication Information Screen**，图 7-2）中，用户可以输入已发表的支持专利申请的科学文献。

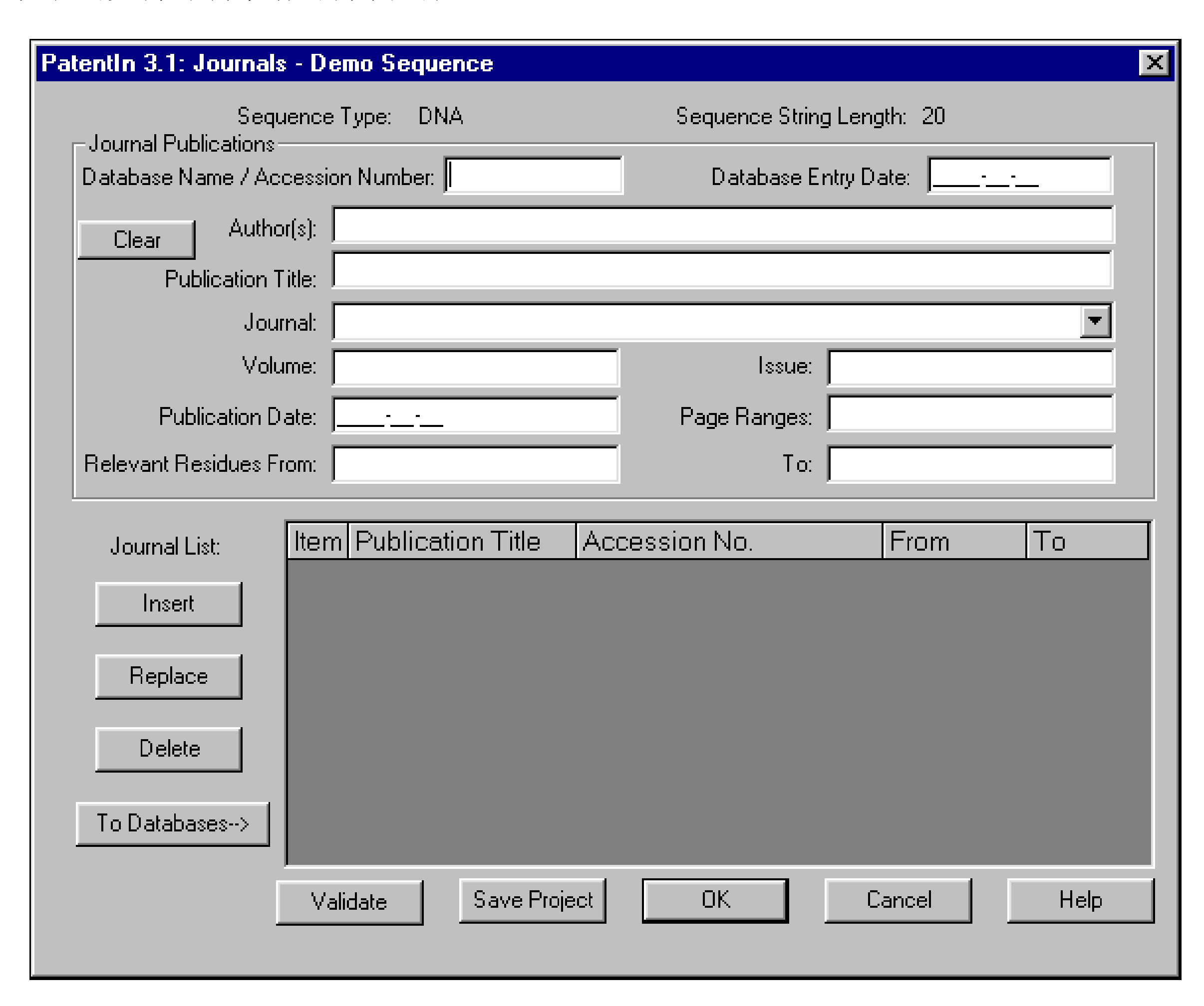


图 **7-2: Journals Publication Information Screen**

输入期刊文献信息：

1. 从申请程序菜单（**Application Steps**），选择文献数据（**Publication Data**），再选择期刊（**Journal**）选项；
2. 输入数据库名称/登记号（**Database Name / Accession Number**）；
3. 输入录入数据库的日期（**Database Entry Date**）；
4. 输入作者名（**Author(s)**）；
5. 输入文章标题（**Publication Title**）；.
6. 从下拉选择列表中，选择期刊名（**Journal**），如果列表中没有，键入期刊名；

7. 输入卷数（**Volume**）；
8. 输入期数（**Issue**）；
9. 输入出版日期（**Publication Date**）；
10. 输入页码范围（**Page Ranges**）；
11. 输入相应残基在序列中起（**Relevant Residues From**）止（**To**）的位置;
12. 点击 **Clear** 键清除屏幕中的期刊文献（**Journal Publications**）部分的内容;
13. 在期刊列表（**Journal List**）中插入一个条目，点击插入键 **Insert**;
14. 要替换期刊列表（**Journal List**）中一个条目，选中它，在窗口上端修改，然后点击 **Replace**;
15. 删除期刊列表（**Journal List**）中一个条目，选中并点击 **Delete** 即可;
16. 点击 **Validate**，使输入的信息生效。已输入（插入）到表中的数据随之确认。仍在期刊文献（**Journal Publications**）编辑栏的信息并未插入，并没有被确认;
17. 点击 **Save Project**，保存信息;
18. 点击 **OK** 键，确认并关闭窗口;
19. 点击 **To Databases→**按钮，继续输入下面的文献、数据库数据（**Publication Data**, **Database**）；

7.3 数据库文献信息

在数据库文献信息窗口（**Databases Publication Information Screen**，图 7–3）中输入已公布的支持专利申请的科学数据库。

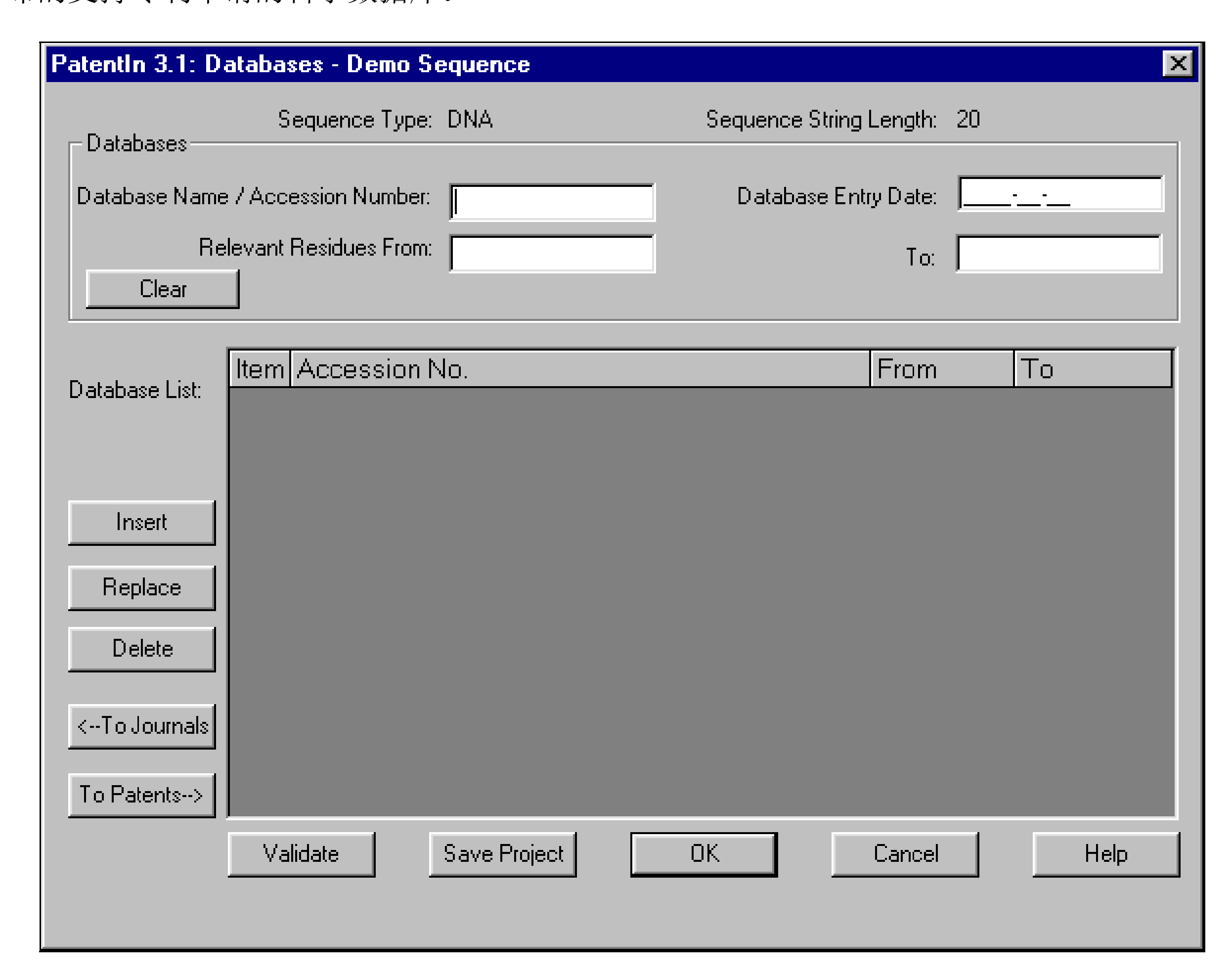


图 7-3: Databases Publication Information Screen

输入数据库文献信息：

1. 从 **Application Steps** 菜单中，选择 **Publications Data**，再选择 **Database**；
2. 输入数据库名称/登记号（**Database Name / Accession Number**）；
3. 输入数据库中的登记日期（**Database Entry Date**）；
4. 输入相应残基在序列中起（**Relevant Residues From**）止（**To**）的位置；
5. 点击 **Clear** 清除窗口中 **Databases** 部分的内容；
6. 在数据库列表（**Database List**）中插入条目，点击 **Insert** 按钮；
7. 替换数据库列表（**Database List**）中的条目，选中要替换的条目，在窗口上端修改后，点击 **Replace**；

8. 删除数据库列表（**Database List**）中的条目，选中它，然后点击 **Delete**；
9. 点击 **Validate**，确认输入的信息。已输入（插入）到表中的数据得到确认。期刊文献编辑栏中的信息并未插入，没有被确认；
10. 点击 **Save Project**，保存信息；
11. 点击 **OK**，确认并关闭窗口；
12. 点击**← To Journals**，返回 **Publication Data**，**Journal** 窗口；
13. 点击 **To Patents→**，继续编辑下一条 **Publication Data**，**Patent** 信息；

7.4 专利文献信息

在专利文献信息窗口（**Patents Publication Information Screen**，图 7-4）中输入已发表的支持专利申请的专利文献。

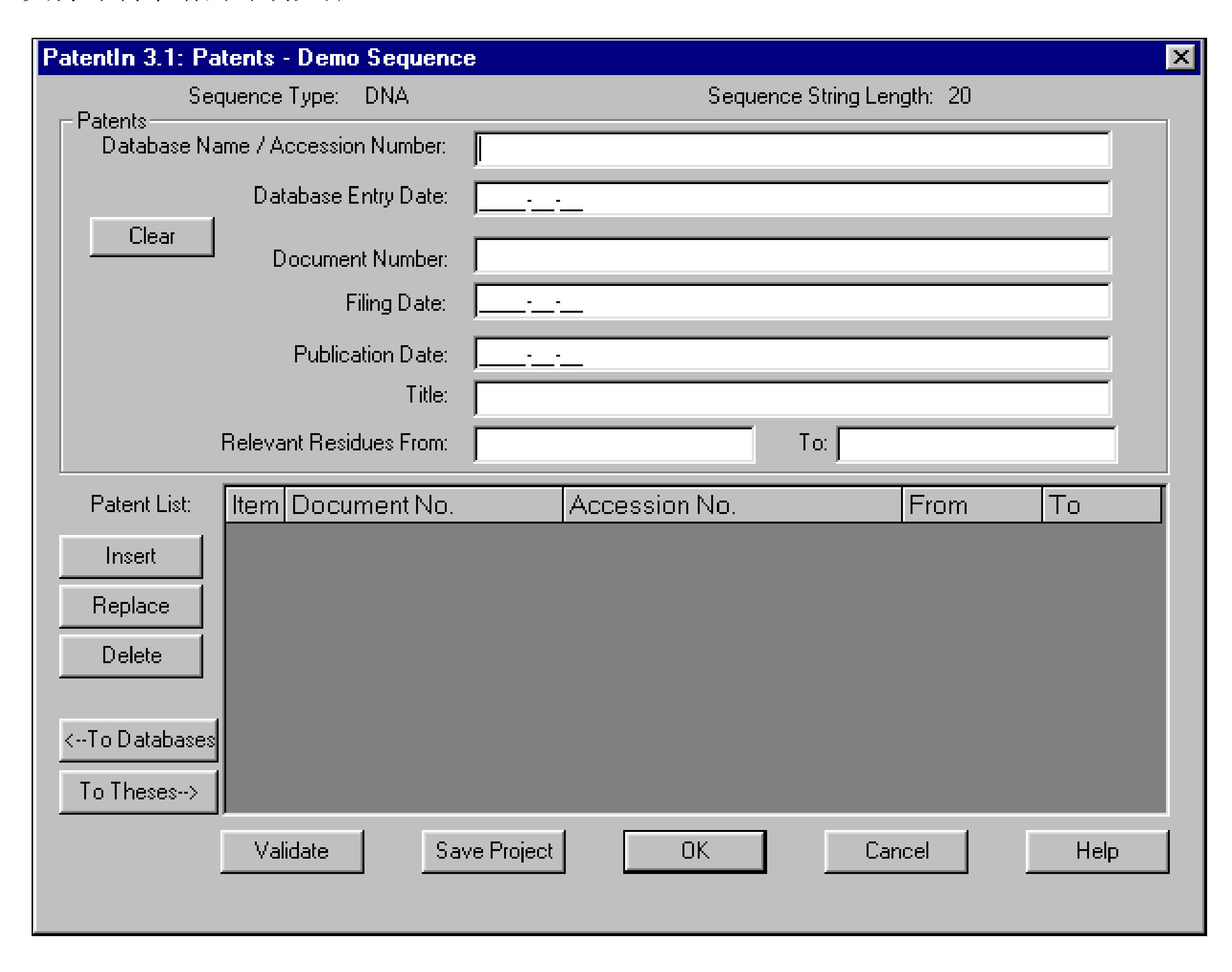


图 7-4: Patents Publication Information Screen

输入专利文献信息：

1. 在 **Application Steps** 菜单中，🖰选择 **Publications Data**，再🖰选择 **Patent**；
2. 🖮输入数据库名称/登记号（**Database Name / Accession Number**）；
3. 🖮输入数据库中的登记日期（**Database Entry Date**）；
4. 🖮输入文献档案号（**Document Number**）；
5. 🖮输入文件归档日期（**Filing Date**）；
6. 🖮输入发表日期（**Publication Date**）；
7. 🖮输入标题（**Title**）；
8. 🖮输入相应残基在序列中的起（**Relevant Residues From**）止（**To**）位置；
9. 🖰点击 **Clear**，清除窗口中 **Patents** 部分的内容；
10. 🖰点击 **Insert**，在专利列表（**Patent List**）中插入一条文献；
11. 在专利列表（**Patent List**）中替换条目，选中要替换的条目，在窗口上端修改，🖰击替 **Replace**；
12. 从专利列表（**Patent List**）删除条目，选中要删除的条目，并点击🖰 **Delete**；
13. 🖰点击 **Validate**，确认输入的信息。已输入（插入）到表中的数据得到确认。期刊文献编辑栏中的信息并未插入，没有被确认；
14. 🖰点击 **Save Project**，保存信息；
15. 🖰点击 **OK**，确认关闭窗口；
16. 🖰点击**← To Databases**，返回前一个 **Publication Data**， **Database** 窗口；
17. 🖰点击 **To Theses→**，继续编辑下一条 **Publication Data**，**Thesis** 信息。

7.5 论文文献信息

论文文献窗口（**Theses Publications Information Screen**，图 7-5），中输入已发表的支持专利申请的论文文献。

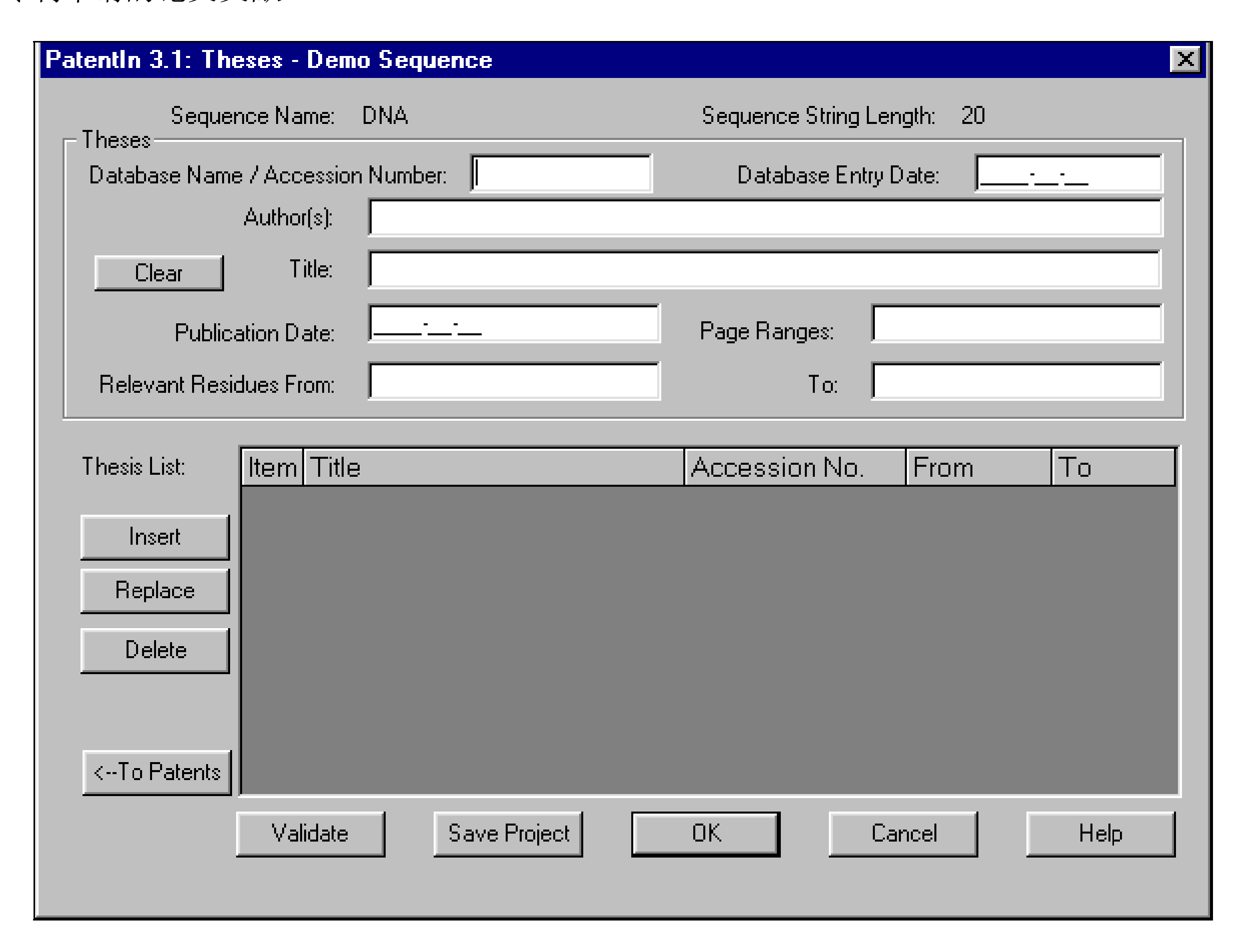


图 **7-5: Theses Publications Information Screen**

输入论文文献信息：

1. 在 **Application Steps** 菜单中，选择 **Publications Data**，再选择 **Thesis**;
2. 输入数据库名称/登记号（**Database Name / Accession Number**）；
3. 输入录入数据库的日期（**Database Entry Date**）；
4. 输入作者名（**Author Name**）；
5. 键入标题（**Title**）；
6. 输入发表日期（**Publication Date**）；
7. 输入页码范围（**Page Ranges**）；
8. 输入相应残基在序列中的起（**Relevant Residues From**）止（**To**）位置；

9. 点击 **Clear**，清除窗口中 **Theses** 部分的内容；
10. 点击 **Insert**，在论文列表（**Thesis List**）中插入一条信息；
11. 替换论文列表（**Thesis List**）中一个条目，选中要替换的条目，在窗口上端修改，点击 **Replace**；
12. 删除论文列表（**Thesis List**）中一个条目，选中要替换的条目，点击 **Delete**；
13. 点击 **Validate**，确认输入的信息。已输入（插入）到表中的数据得到确认。期刊文献编辑栏中的信息并未插入，没有被确认；；
14. 点击 **Save Project**，保存信息；
15. 点击 **OK**，确认并关闭窗口；
16. 点击**←To Patents**，返回前一个 **Publication Data**， **Patent** 窗口。

第八部分
建立任务中的序列列表文件

第八部分
建立任务中的序列列表文件

8.1 序列列表文件

序列列表（sequence listing）文件遵循 ST.25 规范。PatentIn3.x 产生的序列列表，名称为任务名称后加扩展名"ST.25.txt"。这个 ST25 标准文件和任务位于同一目录下。

8.2 建立序列列表文件

序列生成窗口（**Sequence Generation Screen**，图 8-1），提示用户即将生成列表，可以选择出错时给予提示。

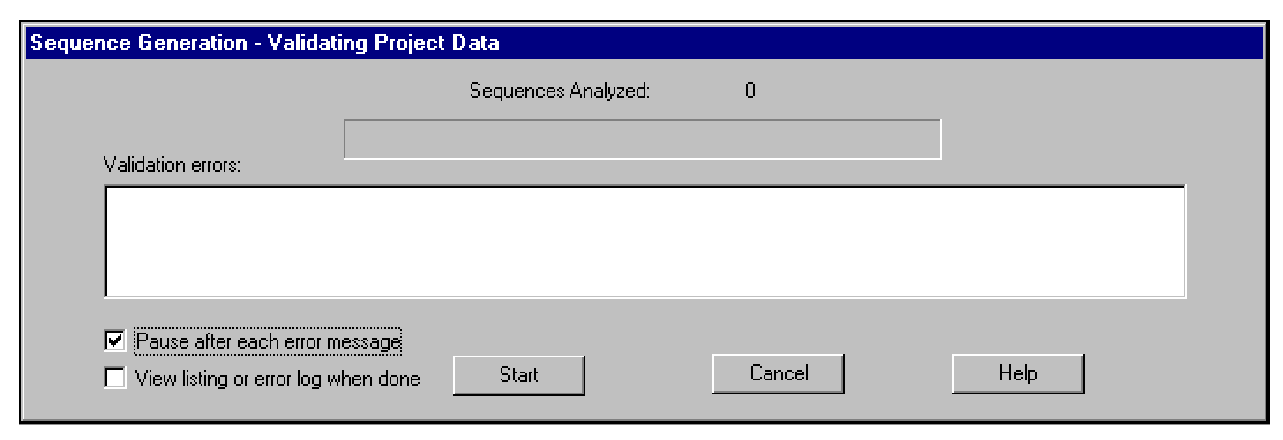


图 8-1: Sequence Generation Screen

生成序列列表：

1. 选择 **Application Steps** 菜单（图 4-1）中的 **Generate Sequence Listing**，或在 **PatentIn** 工具条中选择 **Gen** 选项；
2. 如果希望在序列数据出错时得到提示， 在“Pause after each error message”旁边的方框打勾；
3. 如果希望在生成后直接查看序列列表或错误记录， 在“View listing or error log when done”旁边的方框打勾。生成过程结束时，列表/记录会自动显示；
4. 点击 **Start**，生成序列；

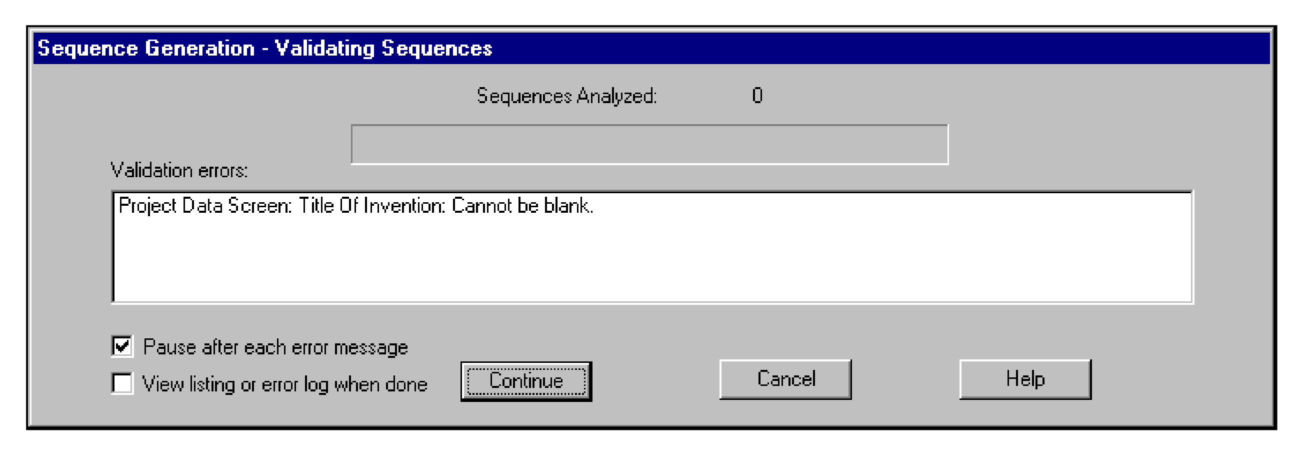


图 **8-2: Second Sequence Generation Screen**

5. 🖰 点击 **Continue** 继续确认（图 8-2）；
6. 🖰 点击 **Cancel** 键取消确认；
7. 如果选择了“Pause after each error message”并由出错信息显示，则会出现出错信息，并暂停确认。

8.3 查看序列列表文件

查看序列列表文件:

1. 如果序列成功生成并选择了“**View listing or error log when done**”，序列列表会自动显示。如果未选择“**View listing or error log when done**”，可以选择 **Project** 菜单中的 **View Sequence Listing** 来查看序列列表;

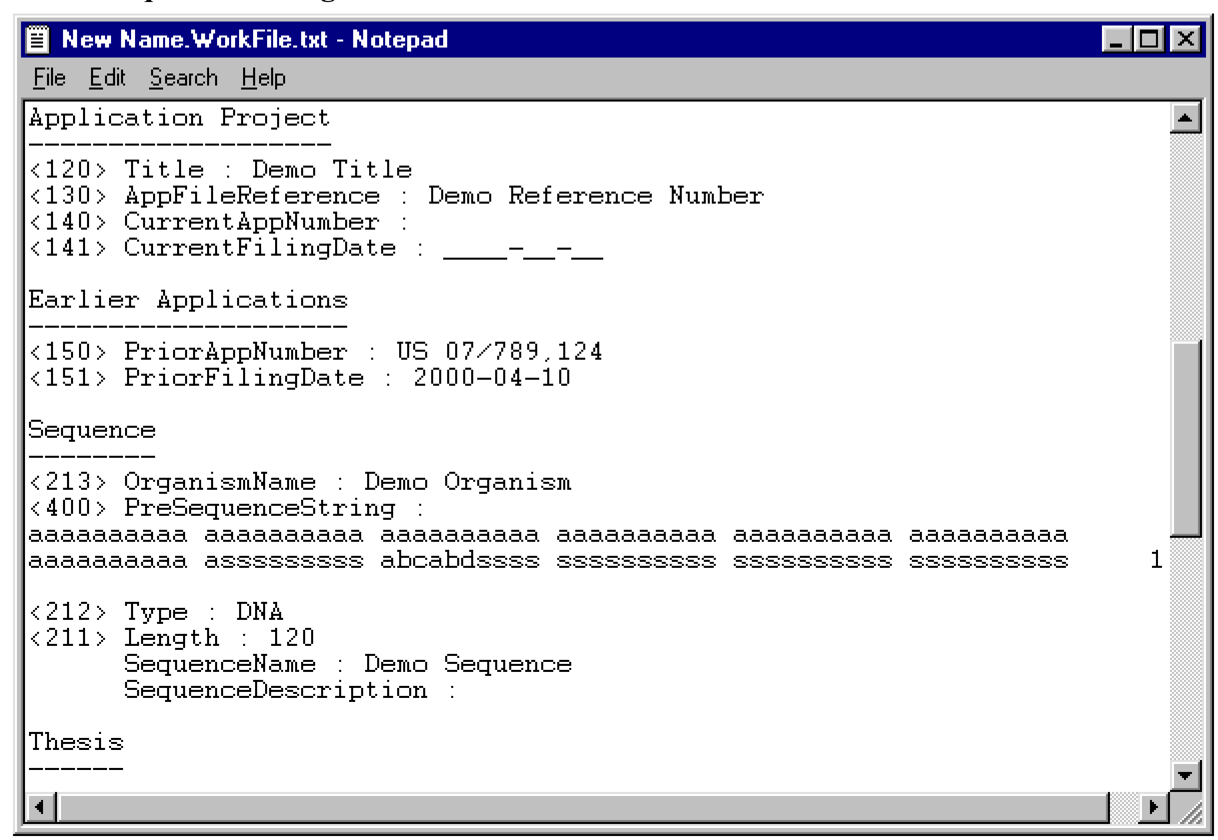


图 **8-1: View Results Window**

ⓘ 对巨大序列或大量序列用户的提示： USPTO 找到一个非常适用于非常大的文本文件的程序，其安装后 60 天内有效的试用版可以从 www.fileviewer.com <http://www.fileviewer.com>下载。程序名称为"V"，版本号为 2000 SR-1。PTO 用 60MB 和 120MB 大的文件进行了测试，结果“非常棒”，也应该可以处理任意大小的文件。PTO 是在膝上型计算机、Win98 环境下进行测试的（安装时可能需要管理员权限）。。美国专利与商标部并非是在推荐使用该产品，只是将其做为有用产品中的一个例子。

8.4 拷贝序列列表到磁盘

拷贝到磁盘窗口（**Copy to Disk Screen**，图 8-1）中，用户可以指定被拷贝文件所在驱动器的名称，拷贝后的文件名称和类型。

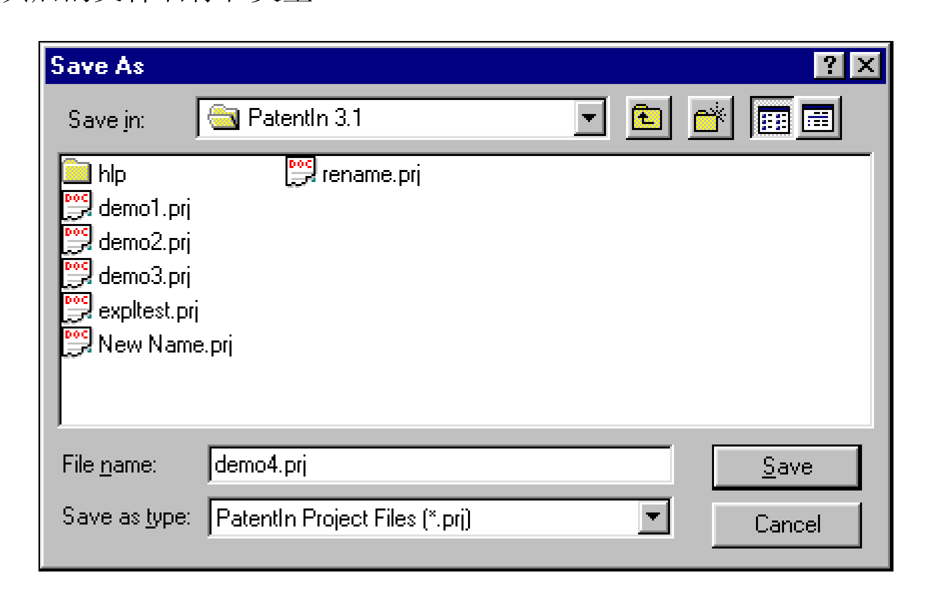


图 8-1: Copy to Disk Screen

拷贝序列列表：

1. 在 **Application Steps** 菜单（图 4–1）中，选择 **Copy to Disk** 选项；
2. 在 **Save in** 栏中，输入驱动器名（图 8-1）；
3. 在 **File name** 栏中，输入文件名；
4. 在 **Save in** 栏中，选择文本（.txt）或压缩（.zip）类型；
5. 点击 **Save to** 进行保存；

如果选择了**.txt** 类型，PatentIn 将检查是否有足够的可用磁盘空间来保存序列列表文件。够的话，文件将被拷贝至指定位置；不够，PatentIn 将建议以压缩格式（**.zip**）保存；

提示: 仅对软盘使用压缩（**.zip**）格式，并在写入前对软盘进行格式化。

提示: 通常，如果选择了硬盘驱动器，将提示用户选择移动存储介质为拷贝目标。

提示: 保存到 CD 命令指 CD 具有其他驱动器一样的功能。也就是说，把 CD 当成一个磁盘驱动器（如 F 驱），用户可以执行浏览类型的命令。如果您的 CD 没有这样的驱动器，文件将存放于其生成的硬盘驱动器上，名称为任务名称后加扩展名.ST25.txt。

附录 **A**
缩略语列表

附录 A
缩略语列表

ARIPO	African Regional Industrial Property Organization （非洲知识产权组织）
ASCII	American Standard Code for Information Interchange（美国信息交换标准码）
CDS	CoDing Sequence（编码序列）
CSC	Computer Sciences Corporation（计算机科学公司）
DLL	Dynamic Link Library （动态链接库文件）
DNA	Deoxyriboneucleic Acid（脱氧核糖核酸）
EPO	European Patent Office (EPO)（欧洲专利局）
FQT	Functional Qualification Test / Functional Quality Testing（功能质量测评）
GPI	Glycosyl-phosphatidylinositol（糖基磷脂酰肌醇）
LTR	Long Terminal Repeat （长末端重复序列）
MB	Megabytes（兆字节）
MHz	Megahertz（兆赫）
PC	Personal Computer（个人电脑）
PCR	Primary Coding Region（一级编码区）
PTO	Patent and Trademark Office（专利与商标管理局）
RNA	Riboneucleic Acid（核糖核酸）
scRNA	small cytoplasmic RNA（胞浆小 RNA）
STS	Sequence Tagged Site（序列定位标签）
TM02	Task Management Plan（计划任务管理程序）
URL	Universal Resource Locator（统一资源定位器）
USPTO	United States Patent and Trademark Office（美国专利和商标局）
UTR	Untranslated Region or Untranscribed Region（非翻译区或非转录区）
WIPO	World Intellectual Property Organization（世界知识产权组织）
WPI	Web PatentIn（网络 PatentIn）
WWW	World Wide Web（万维网）

附录 **B**
栏名称，标志符，大小和类型

附录 B
栏名称，标志符，大小和类型

下表 B-1 包含了所有出现在数据登记界面的域名的完整列表，及其相应的栏标识符，栏长度和栏类型。栏标识符将 PatentIn 数据分为原始数据文件和序列列表文件。

表 **B-1:** 栏名称，标志符，大小和类型

栏标识符	栏名称	栏长度	栏类型 A-代表字母 N- 代表数字数字
N/A	任务名称	8	AN
<110>	申请人姓名	1200	AN
<120>	发明名称	240	AN
<130>	File Reference	60	AN
<140>	Current Application Number	23	AN
<141>	Current Filing Date	8	N
<150>	Earlier Application Number	23	AN
<151>	Earlier Filing Date	8	N
N/A	Sequence File Name	8	AN
<160>	Number of Sequences	5	N
<170>	Software	60	AN
<210>	Information for SEQ ID No.	5	N
<211>	Length	6	N
<212>	Type	3	A
<213>	Organism	60	AN
<220>	Feature	0	B
<221>	Name/key	20	AN
<222>	Location	12	N
<223>	Other Information	260	AN
<300>	Publication Information	0	B
<301>	Authors	120	AN
<302>	Title (of Publication)	120	AN

表 **0-**错误！仅有主文档。**：栏名称，标志符，大小和类型（）**

栏标识符	栏名称	栏长度	栏类型 A-代表字母 N- 代表数字数字
<303>	Journal (name)	40	AN
<304>	Volume	5	AN
<305>	Issue	5	AN
<306>	Pages	20	AN
<307>	Date	30	AN
<308>	Database Accession Number	45	AN
<309>	Database Entry Date	8	N
<310>	Document Number	18	AN
<311>	Filing Date	8	N
<312>	Publication Date	8	N
<313>	Relevant Residues	20	N
<400>	Sequence Description	100,000	AN

附录 **C**
国家代码

附录 C
国家代码

表 C-1 包括完成**任务数据窗口**（**Project Data Screen**，图 4–2）中的 **Current Application Number** 栏和**优先申请信息窗口**（**Prior Application Information Screen**，图 4–3）中的 **Prior Application Number** 栏所用到的国家代码的完整列表。

表 C-1: 国家代码

代码	国家
AF	阿富汉
OA	非洲知识产权组织
AP	非洲工业产权组织
AL	阿尔巴尼亚
DZ	阿尔及利亚
AO	安哥拉
AI	安圭拉岛
AG	安提瓜岛和巴布达岛
AR	阿根廷
AU	澳大利亚
AT	奥地利
BS	巴哈马群岛
BH	巴林
BD	孟加拉国
BB	巴巴多斯岛
BE	比利时
BZ	伯利兹城
BX	荷、比、卢三国联盟商标和外观设计局

表 **C-1:** 国家代码 （续表）

代码	国家
BJ	贝宁湾
BM	百慕大群岛
BT	不丹
BO	玻利维亚
BW	博茨瓦纳
BR	巴西
VG	英属维尔京群岛
BN	文莱达鲁萨兰国
BG	保加利亚
BF	布基纳法索
BU	缅甸
BI	布隆迪
CM	喀麦隆
CA	加拿大
CV	佛得角
KY	开曼群岛
CF	中非共和国
TD	乍得
CL	智利
CN	中国
CO	哥伦比亚
KM	科摩罗
CG	刚果
CR	哥斯达黎加
CI	科特迪瓦
CU	古巴
CY	塞浦路斯

表 **C-1:** 国家代码 （续表）

代码	国家
CS	捷克斯洛伐克
KH	柬埔寨民主共和国
KP	朝鲜人民民主共和国
YD	也门民主共和国
DK	丹麦
DJ	吉布提
DM	多米尼加
DO	多米尼加共和国
EC	厄瓜多尔
EG	埃及
SV	萨尔瓦多
GQ	赤道几内亚
ET	埃塞俄比亚
EP	欧洲专利局
FK	福克兰群岛(马尔维纳斯群岛)
FJ	斐济
FI	芬兰
FR	法国
GA	加蓬
GM	冈比亚
DD	民主德国
DE	联邦德国
GH	加纳
GI	直布罗陀
GR	希腊
GD	格林纳达
GT	危地马拉

表 **C-1:** 国家代码 （续表）

代码	国家
GN	几内亚
GW	几内亚比绍
GY	圭亚那
HT	海地
VA	梵蒂冈
HN	洪都拉斯
HK	香港
HU	匈牙利
IS	冰岛
IN	印度
ID	印尼
IR	伊朗
IQ	伊拉克共和国
IE	爱尔兰
IL	以色列
IT	意大利
JM	牙买加
JP	日本
JO	约旦
KE	肯尼亚
KI	基里巴斯
KW	科威特
LA	老挝国
LB	黎巴嫩
LS	莱索托
LR	利比里亚
LY	利比亚

表 **C-1:** 国家代码 （续表）

代码	国家
LI	列支敦士登
LU	卢森堡公国
MG	马达加斯加岛
MW	马拉维
MY	马来群岛
MV	马尔代夫
ML	马里
MT	马耳他
MR	毛利塔尼亚
MU	毛里求斯
MX	墨西哥
MC	摩纳哥
MN	蒙古
MS	蒙特塞拉特岛
MA	摩洛哥, 摩洛哥皮
MZ	莫桑比克
NR	瑙鲁
NP	尼泊尔
NL	荷兰
AN	荷属安地列斯群岛
NZ	新西兰
NI	尼加拉瓜
NE	尼日尔
NG	尼日利亚
NO	挪威
OM	阿曼
PK	巴基斯坦

表 **C-1:** 国家代码 （续表）

代码	国家
PA	巴拿马
PG	巴布亚新几内亚
PY	巴拉圭
PE	秘鲁
PH	菲律宾共和国
PL	波兰
PT	葡萄牙
QA	卡塔尔
KR	朝鲜共和国
RO	罗马尼亚
RW	卢旺达
KN	圣基茨和尼维斯
SH	圣赫勒拿岛
LC	圣卢西亚岛
VC	圣文森特和格林纳丁斯
WS	萨摩亚群岛
SM	圣马力诺
ST	圣多美和普林西比
SA	沙特阿拉伯
SN	塞内加尔
SC	塞舌尔
SL	塞拉利昂
SG	新加坡
SB	所罗门群岛
SO	索马里
ZA	南非
SU	苏联

表 **C-1:** 国家代码 （续表）

代码	国家
ES	西班牙
LK	斯里兰卡
SD	苏丹
SR	苏里南
SZ	斯威士兰
SE	瑞典
CH	瑞士
SY	叙利亚共和国
TW	台湾
TH	泰国
TG	多哥
TO	汤加
TT	特立尼达和多巴哥
TN	突尼斯
TR	土耳其
TV	图瓦卢
UG	乌干达
AE	阿拉伯联合酋长国
GB	(大不列颠)联合王国
TZ	坦桑尼亚联合共和国
US	美利坚合众国
UY	乌拉圭
VU	瓦努阿图
VE	委内瑞拉
VN	越南
WO	世界知识产权组织
YE	也门

表 **C-1:** 国家代码 （续表）

代码	国家
YU	南斯拉夫
ZR	扎伊尔
ZM	赞比亚
ZW	津巴布韦

附录 **D**
核酸三联体密码和单或三字母氨基酸符号对应表

附录 **D**
核酸三联体密码和单或三字母氨基酸符号对应表

表 D-1 提供了在序列说明域中键入 PRT/1 型数据，输入 PRT/3 型数据和转换 PRT/1 型数据时被接受的符号的清单。在序列说明域转换在生成序列表项目文件时所有的具 CDS 特征的（三核苷酸）密码子被翻译成氨基酸缩写名，此时会用到三核苷酸等价数据。

表 **D-1:** 核酸三联体密码和单或三字母氨基酸符号对应表

PRT/1	PRT/3	对应的核酸
A	Ala	gcu, gcc, gca, gcg, gct
R	Arg	cgu, cgc, cga, cgg, cgt, aga, agg
N	Asn	aau, aac, aat
D	Asp	gau, gac, gat
B	Asx	
C	Cys	ugu, ugc, tgt, tgc
Q	Gln	caa, cag
E	Glu	gaa, gag
Z	Glx	
G	Gly	ggu, ggc, gga, ggg, ggt
H	His	cau, cac, cat
I	Ile	auu, auc, aua, att, atc, ata
L	Leu	uua, uug, cuu, cuc, cua, cug, tta, ttg, ctt, ctc, cta, ctg
K	Lys	aaa, aag
M	Met	aug, atg
F	Phe	uuu, ucc, ttt, ttc
P	Pro	ccu, ccc, cca, ccg, cct
S	Ser	ucu, ucc, uca, ucg, tct, tcc, tca, tcg, agu, agc, agt
T	Thr	acu, acc, aca, acg, act

表 **D-1:** 核酸三联体密码和单或三字母氨基酸符号对应表 **(续)**

PRT/1	**PRT/3**	对应的核酸
W	Trp	ugg, tgg
Y	Tyr	uau, uac, tat, tac
V	Val	guu, guc, gua, gug, gtt, gtc, gta, gtg
X	Xaa	any set containing “n”

附录 **E**
核酸序列特征

附录 E
核酸序列特征

表 E-1（按字母数字顺序）提供了在您事先选择 DNA 或 RNA 作为**序列类型**（**Sequence Type**）并 单击**名称**（**Names**）按钮时，在特征界面中显示的**核苷酸**（**Nucleotide**）序列特征的列表。单击**核苷酸**（**Nucleotide**）方框底部的下箭头后核苷酸列表出现在点选列表中。单击点选列表中的序列特征后，序列特征名称出现在 **Feature Name / Key** 栏中（<221>）。

表 E-1: 核酸序列特征

关键词	说明
allele	等位基因。亲近的个体或菌株间在某位置或其他地方含有某基因稳定的其它形式。
attenuator	衰减子。（1）DNA 上的转录终止调节区域，它控制了一些细菌操纵子的表达；（2）位于启动子和最近的结构基因之间，导致转录的部分终止的序列。
C_region	免疫球蛋白轻和重链的恒定区；T-细胞受体 α，β 和 γ 链，这些序列包括一或多个外显子
CAAT_signal	CAAT 盒;位于真核生物转录单位起始点的 75bp 上游的保守序列，可能参与结合 RNA 聚合酶；共有序列为 GG(C 或 T)CAATCT
CDS	编码序列。与蛋白质中氨基酸序列的核酸的序列（包括终止密码子）；序列特征中包括核酸序列直接翻译为“概念”的氨基酸
conflict	在同一位点或区域，不同独立来源确定的序列有所不同
D-loop	D-环；线粒体 DNA 内的一段单链，有短的 RNA 序列之配对，置换出这一区域的原始配对 DNA 链；也用于描述在 RecA 蛋白质催化的反应中，侵入的单链替代双链 DNA 的一条链的区域
D-segment	免疫球蛋白重链的多变区，以及 T-细胞受体的 β 链
enhancer	增强子。顺式-作用序列，它增强了某些真核生物启动子的作用，并能在任一方向、与启动子相关的任何位置处 （上游或下游）起作用
exon	外显子。编码剪接 mRNA 部分的基因组区域；可以含有 5'UTR、所有 CDS 和 3'UTR
GC_signal	GC 盒。真核生物转录单位起始点上游保守的富含 GC 的区域，以多拷贝或任何方向存在；共有序列=GGGCGG

表 E–1：核酸序列特征 （续）

关键词	说明
gene	具有生物学意义的区域被作为基因，并被指定名称
iDNA	DNA 干扰序列。几种同源重组过程的任何一种能消除的 DNA 序列
intron	内含子。被转录的 DNA 区段，但通过剪接位于其两侧的序列（外显子），即可从转录本内部除去的序列
J_segment	免疫球蛋白轻链和重链，T-细胞受体 α、β 和 γ 链的连接区段
LTR	长末端重复序列。序列两端直接重复的序列，逆转录病毒中很典型
mat_peptide	成熟肽或蛋白质的编码序列；翻译修饰后成熟的或最终的肽或蛋白质产物的编码序列；不包括终止密码子（与相应的 CDS 不同）
misc_binding	不能用其它结合关键词（primer_bind 或 protein_bind）描述的与另一个组成成分共价或非共价结合的核酸中的位点
misc_difference	特征序列与程序中的条目有所不同，不能用其它差异关键词（conflict，unsure，old_sequence，mutation，variation，allele 或 modified_base）表述
misc_feature	不能用其它特征关键词表述的具有生物学意义的区域；新的或稀有特征
misc_recomb	无法用其它重组关键词（iDNA 和 virion）或来源关键词的标志符（/insertion_seq，/transposon，/proviral）描述的，具有断裂和重新连接点的双螺旋 DNA 任何一般性的、位点特异性或复制重组事件
misc_RNA	不能用其他 RNA 关键词（prim_transcript，precursor_RNA，mRNA，5'clip，3'clip，5'UTR，3'UTR，exon，CDS，sig_peptide，transit_peptide，mat_peptide，intron， polyA_site，rRNA，tRNA，scRNA 和 snRNA）定义的任何转录本或 RNA 产物
misc_signal	不能用其他 Signal 关键词（promoter，CAAT_signal，10_signal，RBS，TATA_signal，-35_signal， GC_signal， polyA_signal，enhancer，attenuator，terminator 和 rep_origin）描述的含有控制或改变基因功能或表达信号的区域
misc_structure	不能用其他 Structure 关键词（stem_loop 和 D-loop）描述的任何二级或三级结构或构象
modified_base	被标示的核苷酸是经修饰的核苷酸，并被指示的分子（在 mod_base 标志符中给出）所取代

表 E–1：核酸序列特征 （续）

关键词	说明
mRNA	信使 RNA；包括 5'非翻译区（5'UTR），编码区（CDS，外显子）， 3 '非翻译区（ 3 'UTR）
mutation	突变。某些位置处，亲近品系的 DNA 序列中发生突然、可遗传的变化
N_region	重排的免疫球蛋白区段之间插入的额外核苷酸
old_sequence	某位置处，所表述的序列修改了以前的序列
polyA_signal	polyA 信号。内切核酸酶切断 RNA 转录本，然后发生聚腺苷酸化的位点；共有序列＝AATAAA
polyA_site	polyA 位点。RNA 转录本上转录后发生聚腺苷酸化加上腺嘌呤残基的位点
precursor_RNA	任何不成熟的 RNA 产物；可包括 5'剪切区（5'clip）， 5'非翻译区（5'UTR），编码区（CDS），插入序列（内含子），3'非翻译区（3'UTR）和 3'剪切区(3'clip)
prim_transcript	初级（最初、未加工的）转录本；包括 5 '剪切区（ 5 'clip）, 5 '非翻译区（ 5 'UTR）,编码区（CDS，外显子），插入序列（内含子），3'非翻译区（3'UTR）和 3'剪切区（3'cli p）
primer_bind	复制起始，转录或逆转录的非共价的引物结合位点，包括合成的例如 PCR 引物中的位点
promoter	启动子。RNA 聚合酶的结合以启动转录的 DNA 位点
protein_bind	核酸上蛋白质的非共价结合位点
RBS	核糖体结合位点
repeat_region	含有重复单位的基因组区域
repeat_unit	单一重复元件
rep_origin	复制起点；核酸复制得到两个相同拷贝的起始位点
rRNA	成熟的核糖体 RNA；将氨基酸装配成蛋白质的核糖核蛋白体（核糖体）中的 RNA 成份
S_region	免疫球蛋白重链的开关区；它参与重链 DNA 的重排，导致来自 B 细胞的不同种类的免疫球蛋白类表达

表 **E–1**：核酸序列特征 （续）

关键词	说明
satellite	基本重复单位中很多短的，相同或相关的串联重复序列。大多与基因组的碱基组成比例或其它性质不同，使得它们可以与主体（主带）基因组 DNA 分离开来
scRNA	小的细胞质 RNA；细胞质或（有时）真核细胞核中，任何小的 RNA 分子
sig_peptide	信号肽编码区；分泌蛋白的N－端编码区；此结构域涉及新生肽与膜的结合；前导序列
snRNA	小的核 RNA；任何核中小的 RNA；一些 snRNA 参与剪接或其它 RNA 加工过程
source	序列的生物来源；此关键词是必须填写的；每条序列至少有一个跨越整个序列的来源关键词；每条序列可允许有一个以上的来源关键词
stem_loop	发卡环；单链 RNA 或 DNA 的相邻（反向）互补序列之间的碱基配对形成的双螺旋区域
STS	序列标记位点：基因组上作为界标并能通过 PCR 检测的短的、单拷贝 DNA 序列；测定一系列 STS 的次序即可作出基因组区域的图谱
TATA_signal	TATA 盒；Goldberg-Hogness 盒；真核生物 RNA 聚合酶 II 转录单位起始点上游约 25bp 处保守并富含 AT 的七聚体，可能参与酶的定位以正确地起始转录；共有序列＝TATA（A或T）A（A或T）
terminator	终止子。DNA 上位于转录本的末端或与启动子区域相邻的序列，该序列可导致 RNA 聚合酶转录终止；也可能是阻抑蛋白的结合位点
transit_peptide	转运肽编码序列；核编码的细胞器蛋白质的N端编码区；此结构域参与蛋白质翻译后运送到细胞器中
tRNA	成熟的转移 RNA，参与核酸序列翻译成氨基酸序列的小 RNA 分子（75－85 个碱基长）
unsure	作者不能确定此区域的准确序列
V_region	免疫球蛋白轻链和重链以及T细胞受体的α， β和γ链的可变区；编码可变的氨基末端部分；由 V__segment，D_segment，N_region 和 J_segment 组成
V_segment	免疫球蛋白轻链和重链以及 T 细胞受体 α，β 和 γ 链的可变区；编码大多数可变区（v_region）和前导肽的最后几个氨基酸

表 **E–1**：核酸序列特征 （续）

关键词	说明
variation	变异。有亲缘关系的生物体带有同一基因的稳定突变体（例如 RFLP，多态性等），在某（或其它）位点处与显示的序列不同
3'clip	加工过程中被切下的转录本前体 3'顶端区域
3'UTR	不翻译成蛋白质的成熟转录本的 3 '末端区域（终止密码子之后）
5'clip	加工过程中被切下的转录本前体的 5 '顶端区域
5'UTR	不翻译成蛋白质的成熟转录本的 5 '末端区域（起始密码子之前）
-10_signal	Pribnow 盒；细菌转录单位起点上游约 10bp 处的保守区域，可能参与结合 RNA 聚合酶；共有序列=TATAAT
-35_signal	细菌转录单位起点上游约 35bp 处的保守六聚体；共有序列=TTGACa[]或 TGTTGACA[]

附录 **F**
氨基酸序列特征

附录 F
氨基酸序列特征

表 F-1（按字母数字顺序）提供了在您选择 PRT 作为**序列类型**（**Sequence Type**），并单击名称（**Names**）按钮时，Features Screen中显示的**氨基酸**（**AminoAcids**）序列特征的列表。单击氨基酸方框底部的下拉箭头后，出现氨基酸选取列表。单击选取列表中的序列特征后，序列特征名称出现在**特征名称/关键词**（**Feature Name / Key**）栏中。

表 F-1: 氨基酸序列特征

关键词	说明
ACT_SITE	涉及酶活性的氨基酸
BINDING	化学基团（辅酶，辅基，等）的结合位点；基团的性质在描述栏进行描述
CA_BIND	钙结合区域的范围
CARBOHYD	糖基化位点；碳水化物（已知的话）的性质在描述栏里描述
CHAIN	成熟蛋白质中多肽链的范围
CONFLICT	不同的论文报道的序列不同
DISULFID	二硫键；FROM 和 TO 分别表示通过一个链内二硫键连接的两个残基；如果 FROM 和 TO 是相同的，则是链间二硫键，交联的性质在描述栏里给出
DNA_BIND	DNA 结合区的范围
DOMAIN	序列中感兴趣的区域的范围；其特征在描述栏给出
HELIX	二级结构；螺旋，例如α-螺旋，3_{10}螺旋，或 Pi-螺旋
INIT_MET	序列以起始密码子甲硫氨酸开头
LIPID	共价结合的脂质成分
MAT_PEPTIDE	成熟肽；翻译修饰后的成熟或最终的肽或蛋白质产物
METAL	金属离子结合位点；描述栏里给出金属的性质
MISC_FEATURE	不能用其它特征关键词描述的有生物学意义的区域；新或稀有的特征
MOD_RES	翻译后的残基修饰
MUTAGEN	经实验操作改变的位点

表 F–1：氨基酸序列特征 （续）

关键词	说明
NON_CONS	非连续残基；表示序列中的两个残基不是连续的，它们之间有些末测序的残基
NON_TER	序列末端的残基不是终止残基；假如是位置 1 ，表示此位置不是完整分子的N-末端；如果是最后一个位置，表示此位置不是完整分子的C-末端；此关键词没有描述栏
NP_BIND	核苷酸结合区域；核苷酸特征在描述栏里给出
PEPTIDE	产生的活性肽的范围
PROPEP	前肽的范围
REPEAT	内部重复序列的范围
SIGNAL	信号序列的范围（前肽）
SIMILAR	与另一蛋白质序列相似的范围；另外序列有关的精确资料在描述栏给出
SITE	序列中其它感兴趣的位点
STRAND	二级结构；例如β-链有氢键连接的β-链或分离的β-桥中的残基
THIOETH	硫醚键；FROM 和 TO 表示通过硫醚键连接的两个残基
THIOLEST	硫醇酯键；FROM 和 TO 表示通过硫醇酯键连接的两个残基
TRANSIT	运转肽的范围（线粒体，叶绿体或微体）
TRANSMEM	跨膜区的范围
TURN	二级结构的转角，例如H-键连的转角（ 3 -转角， 4 -转角或 5 -转角）
UNSURE	序列中的不确定处；指序列的不确定区域或作者不确定的列序
VARIANT	作者报道存在序列的变异体
VARSPLIC	描述选择性剪接产生的序列变异体
ZN_FING	锌指区域的范围

附录 G
修饰残基特征数据表

附录 G
修饰残基特征数据表

附录 G 以字母数字顺序给出了在您事先选择 PRT 作为**序列类型**（**Sequence Type**）并单击名称键时，出现在**特征**（）界面中的附加修饰的残基（MOD_RES）的**序列特征列**（**Sequence Features**）表，单击拾取列表中的序列特征后，**MOD_RES** 出现在**特征名称/关键词栏**（**Feature Name / Key**）中(<221>)，第一次以表 G-1 中的 **MOD_RES 在其它信息域中添加 MOD_RES**（**Add the following MOD-RES to the Other Information field**），第二次以表 G-2 中的 **MOD_RES 在其它信息域中添加 MOD_RES**（**Add the following MOD-RES to the Other Information field**）。您可以选择在**其它信息**（**Other Information**）域中添加 **MOD_RES 的一个或两个**（**Add the following MOD-RES to the Other Information field**s），这些数据将出现在其它信息域中(<223>)。

表 G-1: MOD_RES 序列特征的第一数据列表

关键词	说明
(none)	空格（缺省选项）
ACETYLATION	N-末端或其它区域的乙酰化
AMIDATION	酰胺化，通常指成熟活性肽 C-端氨基酸
BLOCKED	未测定的 N-或 C-末端封闭基团
FORMYLATION	N-末端含甲硫氨酸
GAMMA-CARBOXYGLUTAMIC ACID HYDROXYLATION	天冬酰胺，天冬氨酸，脯氨酸或赖氨酸的羟基化
METHYLATION	甲基化，一般指赖氨酸或精氨酸
PHOSPHORYLATION	丝氨酸，苏氨酸，酪氨酸，天冬氨酸或组氨酸的磷酸化
PROPEP	前肽的范围
PYRROLIDONE CARBOXYLIC ACID	N-末端形成内部环内酰胺的谷氨酸
REPEAT	内部序列重复的范围
SULFATATION	巯基化，一般指酪氨酸

表 **G–2: MOD_RES** 序列特征的第一数据列表

符号	含义	
（none）	空格（缺省选项）	Blank space (default option)
Aad	2-氨基脂肪酸	2-Aminoadipic acid
bAad	3-氨基脂肪酸	3-Aminoadipic acid
bAla	β-丙氨酸，β-氨基丙酸	beta-Alanine, beta-Aminopropionic acid
Abu	2-氨基丁酸	2-Aminobutyric acid
4Abu	4-氨基丁酸，哌啶酸	4-Aminobutyric acid, piperidinic acid
Acp	6-氨基已酸	6-Aminocaproic acid
Ahe	2-氨基庚酸	2-Aminoheptanoic acid
Aib	2-氨基异丁酸	2-Aminoisobutyric acid
bAib	3-氨基异丁酸	3-Aminoisobutyric acid
Apm	2-氨基庚二酸	2-Aminopimelic acid
Dbu	2，4 二氨基丁酸	2,4 Diaminobutyric acid
Des	赖氨素	Desmosine
Dpm	2，2'-二氨基庚二酸	2,2’-Diaminopimelic acid
Dpr	2，3-二氨基丙酸	2,3-Diaminopropionic acid
EtGly	N-乙基甘氨酸	N-Ethylglycine
EtAsn	N-乙基天冬氨酸	N-Ethylasparagine
Hyl	羟基赖氨酸	Hydroxylysine
aHyl	别-羟基赖氨酸	allo-Hydroxylysine
3Hyp	3-羟基脯氨酸	3-Hydroxyproline
4Hyp	4-羟基脯氨酸	4-Hydroxyproline
Ide	异赖氨素	Isodesmosine
aIle	别-异亮氨酸	allo-Isoleucine
MeGly	N-甲基甘氨酸，肌氨酸	N-Methylglycine, sarcosine
MeIle	N-甲基异亮氨酸	N-Methylisoleucine

表 **G–2**：（续表）

符号	含义	
MeLys	6-N-甲基赖氨酸	6-N-Methyllysine
MeVal	N-甲基缬氨酸	N-Methylvaline
Nva	正缬氨酸	Norvaline
Nle	正亮氨酸	Norleucine
Orn	鸟氨酸	Ornithine

附录 **H**
含脂类序列的脂类特征

附录 H
含脂类序列的脂类特征

选择 PRT 作为序列类型（**Sequence Type**），并从特征列表中选取 **LIPID** 后，出现表 H-1，按字母顺序给出了其他含脂类序列的特征。单击**序列特征选择列表**（**Sequence Feature Pick List**）中的序列特征（**Sequence Feature**）后，出现 **Add the following LIPID to the Other Information field** 栏，**特征名称/关键词**（**Feature Name / Key**）栏（<221>）出现脂类。选择 **Add the following LIPID to the Other Information field**，数据将出现在其它信息（**Other Information**）栏（<223>）。

表 **H-1:** 其他含脂类的序列特征

关键词	说明
(none)	空格（缺省选项）
MYRISTATE	通过酰胺键与成熟蛋白质的 N-末端甘氨酸残基或内部的赖氨酸残基结合的豆蔻酸基团
PALMITATE	通过硫酯键与半胱氨酸残基或通过酯键与丝氨酸、苏氨酸残基结合的棕榈酸基团
FARNESYL	通过硫酯键与半胱氨酸残基结合的法尼基
GERANYL-GERANYL	通过硫酯键与半胱氨酸残基结合的香叶基-香叶基基团
GPI-ANCHOR	与成熟蛋白质C-末端残基的α-羧基相连的糖基磷脂酰肌醇（GPI）基团
N-ACYL DIGLYCERIDE	原核生物成熟脂蛋白的 N-末端半胱氨酸，具有酰胺键连接的脂肪酸和通过酯键连接两个脂肪酸的甘油基

附录 **I**
序列描述栏允许的符号

附录 I
序列描述栏允许的符号

表 I-1 包括在序列说明栏中键入 PRT/1 型数据，输入 PRT/3 型数据和转换 PRT/1 型数据时可以使用的符号清单。在生成序列表文件时所有的蛋白质数据被翻译为氨基酸缩略语（PTR/3）数据时会使用 PRT/1 和 PRT/3 数据表。

表 I-1: 序列描述栏允许的符号

DNA	RNA	DNA/RNA	Protein/1	Protein/3
a	a	a	A	Ala
g	g	g	C	Cys
c	c	c	D	Asp
t		t	E	Glu
	u	u	F	Phe
r	r	r	G	Gly
y	y	y	H	His
m	m	m	I	Ile
k	k	k	K	Lys
s	s	s	L	Leu
w	w	w	M	Met
b	b	b	N	Asn
d	d	d	P	Pro
h	h	h	Q	Gln
v	v	v	R	Arg
n	n	n	S	Ser
			T	Thr
			V	Val
			W	Trp
			Y	Tyr
			B	Asx
			Z	Glx
			X	Xaa

附录 **J**
含修饰碱基序列的其他特征

附录 J
含修饰碱基序列的其他特征

表 J-1 以字母数字顺序给出了在您事先选择 DNA 或 RNA 作为**序列类型（Sequence Type）**并单击名称键时，出现在特征界面中的附加**修饰碱基（modified_base）**序列特征的列表。单击选择列表中的序列特征后，**modified_base** 出现在**特征名称/关键词（Feature Name / Key）**栏中(<221>)，**在其它信息栏中添加 modified_base 栏（Add the following modified_base to the Other Information）**出现在特征界面中。用户可从选择列表中选择，数据出现在**其它信息（Other Information）**栏中(<223>)。

表 J-1: 含修饰碱基序列的修饰碱基特征

符号	含义	
ac4c	4-乙酰胞苷	4-acetylcytidine
chm5u	5-（羧羟甲基）尿苷	5-(carboxyhydroxymethyl)uridine
cm	2'-O-甲基胞苷	2'-O-methylcytidine
cmnm5s2u	5-羧甲基氨甲基-2-硫代尿苷	5-carboxymethylaminomethyl-2-thiouridine
cmnm5u	5-羧甲基氨甲基尿苷	5-carboxymethylaminomethyluridine
d	二氢尿苷	dihydrouridine
fm	2'-O-甲基假尿苷	2'-O-methylpseudouridine
gal q	β，D-半乳糖 Q 核苷	beta, D-galactosylqueuosine
gm	2'-O-甲基鸟苷	2'-O-methylguanosine
I	肌苷	inosine
i6a	N6-异戊烯基腺苷	N6-isopentenyladenosine
m1a	1-甲基腺苷	1-methyladenosine
m1f	1-甲基假尿苷	1-methylpseudouridine
m1g	1-甲基鸟苷	1-methylguanosine
m1I	1-甲基肌苷	1-methylinosine
m22g	2'2-二甲基鸟苷	2,2-dimethylguanosine
m2a	2-甲基腺苷	2-methyladenosine

表 **J–1**：含修饰碱基序列的修饰碱基特征（续）

符号	含义	
m2g	2-甲基鸟苷	2-methylguanosine
m3c	3-甲基胞苷	3-methylcytidine
m5c	5-甲基胞苷	5-methylcytidine
m6a	N6-甲基腺苷	N6-methyladenosine
m7g	7-甲基鸟苷	7-methylguanosine
mam5u	5-甲基氨基甲基尿苷	5-methylaminomethyluridine
mam5s2u	5-甲氧基氨基甲基-2-硫代尿苷	5-methoxyaminomethyl-2-thiouridine
man q	β，D-甘露糖 Q 核苷	beta, D-mannosylqueuosine
mcm5s2u	5-甲氧基羰基甲基-2-硫代尿苷	5-methoxycarbonylmethyl-2-thiouridine
mcm5u	5-甲氧基羰基甲基尿苷	5-methoxycarbonylmethyluridine
mo5u	5-甲氧基尿苷	5-methoxyuridine
ms2i6a	2-硫代甲基-N6-异戊烯基腺苷	2-methylthio-N6-isopentenyladenosine
ms2t6a	N-（（9-β-D-呋喃核糖基-2-硫代甲基嘌呤-6-yl）氨基甲酰）苏氨酸	N-((9-beta-D-ribofuranosyl-2-methylthiopurine-6-yl)carbamoyl)threonine
mt6a	N-（（9-β-D-呋喃核糖嘌呤-6-yl）N-甲基氨基甲酰）苏氨酸	N-((9-beta-D-ribofuranosylpurine-6-yl)N-methylcarbamoyl)threonine
mv	尿苷-5-氧化乙酸-甲基酯	uridine-5-oxyacetic acid-methylester
o5u	尿苷-5-氧化乙酸	uridine-5-oxyacetic acid
osyw	wybutoxosine	wybutoxosine
p	假尿苷	pseudouridine
q	Q 核苷	queuosine
s2c	2-硫代胞苷	2-thiocytidine
s2t	5-甲基-2 硫代尿苷	5-methyl-2-thiouridine

表 **J–1**：含修饰碱基序列的修饰碱基特征（续）

符号	含义	
s2u	2-硫代尿苷	2-thiouridine
s4u	4-硫代尿苷	4-thiouridine
t	5-甲基尿苷	5-methyluridine
t6a	N-（（9-β-D-呋喃核糖嘌呤-6-yl）-氨基甲酰）苏氨酸	N-((9-beta-D-ribofuranosylpurine-6-yl)-carbamoyl)threonine
tm	2'-O-甲基-5-甲基尿苷	2'-O-methyl-5-methyluridine
um	2'-O-甲基尿苷	2'-O-methyluridine
yw	Wybutosine	wybutosine
x	3-（3-氨基-3-羧基-丙基）尿苷，（acp3）u	3-(3-amino-3-carboxy-propyl)uridine, (acp3)u

April 27, 2001
Version 2.1
Task Order Number CSC-01-11

附录 **K**
技术提示

附录 K
技术提示

K.1 MICROSOFT® ACCESS 提示

PatentIn 3.1 与含有 PantentIn 2.1 使用的标准数据库文件连接的 Microsoft® Access 的一个程序 patin2xconvert.mdb 一起打包。为使用户查看目录，必须安装连接表管理器。同样的，数据存取文件必须包括 dBASE 5 才能读 PatentIn 2.1 文件。PatentIn 3.1 is packaged with a Microsoft® Access program, patin2xconvert.mdb, that contains links to the standard database files that PatentIn 2.1 uses. In order to enable the user to view the directory, the Link 表 Manager must be installed. Similarly, the Data Access files must include dBASE 5 in order to read the PatentIn 2.1 files.

K.1.1 安装 Microsoft® Access 97

对 Office 97 的一般说明：

1. 运行安装光盘；
2. 选择自定义安装。典型安装不能安装所有必须的组件；
3. 选择 Microsoft® Access；
4. 单击更改选项键；
5. 选择高级安装向导；
6. 返回主界面；
7. 选择数据存取；
8. 单击更改选项键；
9. 选择数据库驱动器；
10. 单击更改选项键；
11. 选择 dBase & Microsoft® FoxPro Drivers.
12. 返回主界面；
13. 结束安装（或添加安装）。

K.1.2 提示与技巧

1. Patin2xconvert.mdb 必须与 CWPI.exe 在同一目录下，且必须包含帮助目录；
2. 每次输入时，Patin2xconvert.mdb 会更新（升级）其连接；
3. Patin2xconvert.mdb 可独立打开，以证实所有 Microsoft® Access 组件都是可用的；

a. 连接表管理器可以在 Office 97 的 Tools | Add-Ins 和 Office 2000 的 Database Utilities 中找到；
b. 在选择文件|获取扩展数据|连接表……时可以检验驱动器，From File Types 文件类型包括 dBASE 5 (*.dbf)。

4. patin2xconvert.mdb 不支持关系式；.
5. patin2xconvert.mdb 不运行代码或查询；
6. 尽管有时必要，但不推荐使用源于 Patin2xconvert.mdb 的分析数据，因为这可能使已连接的数据升级（如果有新数据，确保备份原 Exploring data from patin2xconvert.mdb, though sometimes necessary, is not recommended since it is possible to update the linked data itself. (Be sure the data is backed-up if this is new to you.)
7. MDAC，Microsoft?的可分配组件，可能不足以运行这一输入 MDAC, a distribu 表 component from Microsoft®, may not be sufficient to run this import.

K.2 一般提示与技巧

1. 每台机器安装需要对几个动态链接库文件进行注册（Dynamic Link Library，DLL）；
2. PatentIn 3.x 支持长文件名；
3. 为了缩短处理极长序列的时间，粘贴过程中不能预览（确认）数据；
4. 文件提交命令，磁盘拷贝，在外存储器中备份您硬盘驱动器中已有的文件。通常这不能拷贝到您的硬盘驱动器中 The file submission command, Copy to Disk, copies a file already on your hard drive to an external medium. In general, it cannot be used to copy to your hard drive.
5. 帮助文件是 ASCII 文本文件，可以在本地进行升级，包括被翻译成本国语言；
6. 不出现界面或闪烁通常是没有安装在本机上引起的故障；
7. Windows 2000 下默认为不按 Alt-key 时不显示工具条的下划线。

虽然尚无功能加强版，维持版是必须的，而未报告的问题不能解决。但可以说，未来增强版任何功能的增强不会包括"零售"版 Windows 95，而 Windows 95SP2 版仍将支持。

K.3 安装与测试备忘录

安装 PatentIn 3.1 后，有 3 个新的 DLL 文件被安装在 PatentIn 3.1 目录下，通常是 C:\Program Files\USPTO\PatentIn 3.1：

DDAO35.DLL

MFC42.DLL

MSVCRT.DLL

每个都自动加入注册表，As with any new installation it is recommended that a backup of the system be made first. Should the system network staff wish to unregister these DLLs because of perceived conflicts the commands 建议在任何新安装之前首先进行系统备份。如果由于先前的冲突命令，系统网络员不希望注册 DLLs，可以在开始|运行菜单中执行：

regsvr32 –u “C:\Program Files\USPTO\PatentIn 3.1\ddao35.dll”

regsvr32 –u “C:\Program Files\USPTO\PatentIn 3.1\mfc42.dll”

regsvr32 –u “C:\Program Files\USPTO\PatentIn 3.1\msvcrt.dll”

may be executed from Start|Run. (It is the “–u” that causes these statements to be unregister statements.) For the user, PatentIn is designed to be totally uninstalled from the Control Panel Add/Remove processes（是"-u"使其成为非注册命令）。对用户来说，PatentIn 可用在控制面板中添加或删除来完全卸载。.

K.3.1 配置测试：

进行测试时用到以下配置：

Microsoft® Windows 95

Version 4.00.950 B

Internet Explorer: 3.0(4.70.1158)

Microsoft® Windows 98

Version 4.10.2222 A

Internet Explorer 5.00.2614.3500

Microsoft® Windows NT

Version 4.00,1381

Internet Explorer 5.50.4134.0600

Microsoft® Windows 2000 Professional

Version 5.00.2195

Internet Explorer 5.00.3103.1000

Microsoft® Windows ME

虽然未在这个版本下进行正式测试，但我们的一位 beta 版试用者已告诉我们，没有在该版本中发现问题。

虽然无法详尽测试所有的配置，但相信 PatentIn 不会对以上配置造成任何负面影响。

K.3.2 Internet Explorer 注意事项

Internet Explorer 中有几个 DLL 文件被 Checker 和/或 PatentIn 使用，虽然 PatentIn 已经试图单独使用自己的 DLL 文件，但是已有 Checker/PatentIn 3.0 用户通知 USPTO 其中之一或这两个在 Microsoft® Windows 98 上的 Internet Explorer 4.0 上出现问题。用户迅速发现了问题，并告知 USPTO 升级到 Internet Explorer 5.0 可以解决问题。